TU WIEN TECHNISCHE UNIVERSITÄT WIEN

D I S S E R T A T I O N

Numerical Analysis and Efficient Simulation of Micromagnetic Phenomena

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften unter der Leitung von

Univ.-Prof. Dr. Dirk Praetorius

E101 – Institut für Analysis und Scientific Computing, TU Wien

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Carl-Martin Pfeiler

Matrikelnummer: 01027222

Diese Dissertation haben begutachtet:

1. **Prof. Dr. Dirk Praetorius**
   Institut für Analysis und Scientific Computing, TU Wien

2. **Prof. Dr. Sören Bartels**
   Abteilung für Angewandte Mathematik, Albert-Ludwigs-Universität Freiburg

3. **Prof. Dr. Norbert J. Mauser**
   Fakultät für Mathematik, Universität Wien

Wien, am 10. Februar 2022

Kurzfassung

Zeitabhängige mikromagnetische Phänomene werden üblicherweise durch die nichtlineare Landau–Lifshitz–Gilbert Gleichung (LLG) beschrieben. Die Dynamik von LLG wird vom effektiven Feld induziert, welches vom mikromagnetischen Energiefunktional abgeleitet wird, meist gekoppelt an andere partielle Differentialgleichungen wie dem Maxwell-System.

Diese Dissertation behandelt numerische Herausforderungen der zuverlässigen und effizienten Integration von LLG. Der Fokus liegt auf drei Familien von numerischen Finite-Elemente-Verfahren, welche (unbedingt) gegen eine schwache Lösung von LLG konvergieren: das Midpoint-Verfahren von Bartels & Prohl (2006), das Tangent-Plane-Verfahren von Alouges (2008), und zwei Prädiktor-Korrektor-Verfahren von Kim & Wilkening (2018).

Wir erweitern das Tangent-Plane-Verfahren und das Midpoint-Verfahren auf allgemeinere Energiebeiträge, sodass die nicht standardmäßige Dzyaloshinskii–Moriya Wechselwirkungsenergie (DMI) abgedeckt wird, welche essenziell für die Enukleation und die Stabilisierung von chiralen magnetischen Skyrmionen ist. Unsere Konvergenzanalysis ist konstruktiv und zeigt die Existenz von schwachen Lösungen für LLG mit DMI, was — soweit wir wissen — in der Literatur fehlte. Unsere numerischen Experimente deuten darauf hin, dass in den meisten Szenarien das günstigere Tangent-Plane-Verfahren hinreichend genaue Simulationen ermöglicht. In einer vergleichenden numerischen Studie beobachten wir jedoch, dass für sehr sensitive Dynamiken, welche wesentlich von einer akkuraten Energieevolution abhängen, das Midpoint-Verfahren die zuverlässigsten Resultate liefert.

Wir formulieren und analysieren effiziente Strategien für die Lösung diskreter Systeme, welche im Tangent-Plane-Verfahren oder im Midpoint-Verfahren vorkommen: Wir führen die — soweit wir wissen — erste rigorose Analysis für das nichtlineare Midpoint-Verfahren für dreidimensionalen Mikromagnetismus mit einer Newton-Linearisierung durch. Des Weiteren leiten wir für das nicht-symmetrische lineare Tangent-Plane-System, welches im zeitabhängigen diskreten Tangentialraum gestellt ist, mögliche Vorkonditionierungsstrategien her, die zu linearer Konvergenz des vorkonditionierten GMRES-Algorithmus führen.

Für die Prädiktor-Korrektor-Verfahren von Kim & Wilkening schließen wir eine fundamentale analytische Lücke indem wir deren unbedingte Wohldefiniertheit zeigen. Unsere Analysis deckt selbst den Fall eines verschwindenden Gilbert-Dämpfungsparameters ab. Zusätzlich schlagen wir implizit-explizite Adaptionen der Verfahren vor, welche den Rechenaufwand drastisch reduzieren und dabei deren formale Konvergenzordnung erhalten.

Unsere theoretischen Beiträge werden stets durch numerische Experimente belegt und weitergehend untersucht. Ein zentraler Punkt dieser Dissertation ist die Entwicklung eines leicht zu benutzenden *open-source* Softwaremoduls für die Simulation von mikromagnetischen Phänomenen: Das Modul *Commics* basiert auf der Finite-Elemente-Bibliothek NGSolve, ist frei verfügbar auf GitLab, implementiert die in dieser Dissertation untersuchten Finite-Elemente-Integratoren, und stellt ein Werkzeug zur Verfügung, um die Forschung zur numerischen Integration von LLG zu verbessern und zu fördern.

Abstract

Time-dependent micromagnetic phenomena are usually described by the nonlinear Landau–Lifshitz–Gilbert equation (LLG). The driving force of LLG is the effective field, which is derived from the micromagnetic energy functional, usually coupled to other partial differential equations like the Maxwell system.

This thesis addresses some of the numerical challenges of the reliable and efficient integration of LLG. The focus is on three families of finite element-based numerical schemes that are proven to be (unconditionally) convergent towards a weak solution of the problem: the midpoint scheme by Bartels & Prohl (2006), the tangent plane scheme by Alouges (2008), and two recent predictor-corrector methods by Kim & Wilkening (2018).

We extend the tangent plane scheme and the midpoint scheme to more general energy contributions, covering the non-standard Dzyaloshinskii–Moriya interaction (DMI) energy, which is the essential ingredient for the enucleation and the stabilization of chiral magnetic skyrmions. Our constructive convergence analysis proves the existence of weak solutions in presence of DMI, which — to the best of our knowledge — was missing in the literature. Our numerical experiments hint that in most scenarios the cheaper tangent plane scheme provides sufficiently accurate simulations. In a comparative numerical study we observe, however, that for very sensitive dynamics crucially relying on an accurate energy evolution the midpoint scheme yields the most reliable results.

We propose and analyze strategies for the efficient solution of discrete systems obtained from the tangent plane scheme or from the midpoint scheme: We provide — to the best of our knowledge — the first rigorous analysis of the nonlinear midpoint scheme for three dimensional micromagnetics linearized by Newton's method. Further, for the constrained and non-symmetric linear system arising from the tangent plane scheme and posed in the time-dependent discrete tangent space, we derive possible preconditioning strategies, which guarantee linear convergence of the preconditioned GMRES algorithm.

For the predictor-corrector methods recently proposed by Kim & Wilkening, we close a fundamental gap in the original work by establishing unconditional well-posedness of the schemes. Our analysis even covers the case of a vanishing Gilbert damping parameter. Moreover, we propose implicit-explicit versions of the predictor-corrector methods, drastically reducing the computational cost, while preserving the formal convergence order.

All our theoretical contributions are accompanied by supportive numerical experiments. As a central element of the thesis, inspired by our theoretical findings, an easy-to-use open-source software module for the simulation of micromagnetic phenomena is developed: The module *Commics* is based on the multiphysics finite element library NGSolve, is freely available on GitLab, implements the state-of-the-art finite element integrators discussed in this thesis, and thus provides a means to enhance and promote research on the numerical integration of LLG.

Danksagung

In erster Linie gilt mein Dank Professor Dirk Praetorius für die Möglichkeit mein Doktorats-studium als Teil seines Teams am Institut für Analysis und Scientific Computing absolvieren zu dürfen: Neben der fachlichen Expertise und Deiner ständigen Verfügbarkeit will ich ins-besondere den wissenschaftlichen Freiraum erwähnen, den Du mir stets gegeben hast — das ist nicht selbstverständlich. Danke für Dein Vertrauen, Dirk!

Großer Dank gilt Professor Sören Bartels und Professor Norbert J. Mauser für die gewis-senhafte Begutachtung dieser Dissertation.

Ich danke meinen Kollegen Alex H., Alex R., Andrea, Ani, Bernhard, Giovanni, Gregor, Hubert, Julian, Markus, Max Be., Max Br., Michael F., Michael I., Michele, Simon, Stefan und Ursula, dass ich mich in der Arbeitsgruppe sowohl fachlich als auch sozial stets in bester Gesellschaft wissen konnte. An dieser Stelle danke ich insbesondere Michele Ruggeri für sein Co-Mentoring und die hervorragende Zusammenarbeit all die Jahre über. Außerdem danke ich Dirk Pauly dafür, dass er mir ungeachtet widrigster Bedingungen 2020 einen tollen Forschungsaufenthalt an der Universität Duisburg-Essen ermöglicht hat. In diesem Zusammenhang richte ich auch ein herzliches Dankeschön an Professor Johannes Kraus.

Abseits der Arbeit danke ich meinen studentischen Weggefährten und Freunden Chri-stoph, Daniel, Gerold, Gregor, Patrick und Steffi für die unzähligen geselligen Abende. Ich bedanke mich für den sportlichen Ausgleich bei Aldin, Andi, Chyster, Daniel, Kushtrim und Will, sowie allen anderen die regelmäßig, auch nachts und im Winter, mit mir im Erlachpark zocken. Danke Anna, Dino, Edi, Ferdi, Jan, Jay, Lena, Leo, Marius, Max, Mona, Nedim und Paul für all die unvergesslichen Erinnerungen, die nach bald 20-jähriger Freundschaft immer wieder ergänzt werden.

Zu guter Letzt danke ich meiner Familie: Danke euch allen, dass ich mich immer wieder auf den nächsten Heimatbesuch freue. Danke für eure Unterstützung. Danke, dass ich mich immer auf euch verlassen kann. Danke Francesca, dass du an mich glaubst, mich immer unterstützt und deine Zeit mit mir teilst.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 10. Februar 2022

_____
Carl-Martin Pfeiler

Contents

List of Figures

List of Tables

1 Introduction

This thesis is concerned with the efficient numerical simulation of magnetization dynamics in ferromagnets at sub-micrometer length scales, which is encountered in many technological applications, like, e.g., magneto-resistive storage devices (HDD, MRAM), magnetic sensors, and recording heads. At this length scale, ranging from few nanometers to micrometers, singular magnetic structures like domain walls and vortices are resolved. At the same time, the scale allows to average the atomic structure of a ferromagnet. A well-accepted description of this so-called *continuum approximation* or *continuous medium approximation* is the nonlinear *Landau–Lifshitz–Gilbert* equation (LLG). Apart from the theoretical understanding of magnetization dynamics, numerical integration of LLG plays an essential role in the investigation and prediction of micromagnetic phenomena and hence in the improvement of existing and the development of future magnetic devices.

1.1 On dynamical micromagnetism

A three-dimensional micromagnetic configuration inside some bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$ is mathematically described by a vector field $\boldsymbol{M} \colon \Omega \to \mathbb{R}^3$. If the temperature is assumed to be constant inside the ferromagnetic medium Ω and far below the *Curie temperature*, then the magnetization length is constrained to be uniform $|\boldsymbol{M}| \equiv M_s > 0$ over all of Ω, where the constant M_{s} is the so-called *saturation magnetization*. After normalization $\boldsymbol{m} := \boldsymbol{M}/M_s$, the micromagnetic state is solely described by the function $\boldsymbol{m} \colon \Omega \to \mathbb{S}^2 := \{\boldsymbol{x} \in \mathbb{R}^3 \colon |\boldsymbol{x}| = 1\}$, which we call *magnetization*; see Section 1.3.1 below for details on the nondimensionalization of the physical problem. Any such magnetization state can be assigned a value $\mathcal{E}(\boldsymbol{m}) \in \mathbb{R}$ representing its micromagnetic energy, which comprises several energy contributions specified in Section 1.2 below.

The most fundamental contribution to the micromagnetic energy functional is the so-called *exchange (energy) contribution*

$$\mathcal{E}_{\mathrm{ex}}(\boldsymbol{m}) := A \int_{\Omega} |\boldsymbol{\nabla}\boldsymbol{m}|^2 \, \mathrm{d}\boldsymbol{x}\,, \tag{1.1}$$

with the material-dependent *exchange stiffness constant A* (in J/m). This exchange energy contribution models that neighbouring magnetic moments tend to align locally with each other.

If the micromagnetic configuration $\boldsymbol{m}$ is not at equilibrium, a configuration (locally) minimizing the micromagnetic energy $\mathcal{E}(\boldsymbol{m})$, then over time the magnetization is driven towards such a (locally) minimal energetic state.

A well-accepted description of the magnetization dynamics up to time $T > 0$ is the phenomenological Landau–Lifshitz–Gilbert equation (LLG) [Gil55, LL08], which in Landau–

Lifshitz form reads

$$(1 + \alpha^2)\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) - \alpha \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})) \qquad \text{in } (0, T) \times \Omega. \qquad (\mathcal{LL})$$

Here, $0 < \alpha \leq 1$ denotes the *Gilbert damping parameter*. The so-called effective field $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})$ is given by the (rescaled) negative functional derivative of the system's micromagnetic energy

$$\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) := -\frac{1}{\mu_0 M_{\mathrm{s}}^2} \frac{\delta \mathcal{E}(\boldsymbol{m})}{\delta \boldsymbol{m}}, \qquad (1.2)$$

where $\mu_0 = 4\pi \cdot 10^{-7}\,\mathrm{N/A}^2$ denotes the *vacuum permeability*. The system $(\mathcal{LL})$ is endowed with an initial condition

$$\boldsymbol{m}(0) = \boldsymbol{m}^0 \colon \Omega \to \mathbb{S}^2 \qquad \text{in } \Omega, \qquad (1.3)$$

and natural boundary conditions that, in the exchange-only case $\mathcal{E}(\boldsymbol{m}) = \mathcal{E}_{\mathrm{ex}}(\boldsymbol{m})$, turn out to be homogeneous Neumann boundary conditions on the parabolic boundary

$$\partial_{\boldsymbol{n}} \boldsymbol{m} = \boldsymbol{0} \qquad \text{on } (0, T) \times \partial\Omega, \qquad (1.4)$$

where $\boldsymbol{n} \colon \partial\Omega \to \mathbb{R}^3$ denotes the outward-pointing unit normal vector on $\partial\Omega$. In the literature, this setting, where all energy contributions but the exchange energy are neglected, is referred to as *small particle limit* of LLG. In the mathematics community, this simplified model is popular as it allows for a clear presentation of theory and new results. In particular, introducing the so-called *exchange-length* $\ell_{\mathrm{ex}} := \sqrt{2A/(\mu_0 M_{\mathrm{s}}^2)}$ (measured in m), by (1.1) and (1.2) the exchange energy yields the highest-order (i.e., second-order) contribution $\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m}$ of the effective field $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})$, and hence, this simplified model covers the characteristic nature of LLG. Consequently, any micromagnetic model must cover at least this fundamental energy contribution $\mathcal{E}_{\mathrm{ex}}(\boldsymbol{m})$.

The sole structure of $(\mathcal{LL})$ gives insight on the evolution dynamics: Taking the scalar product of $(\mathcal{LL})$ with the magnetization $\boldsymbol{m}$ reveals the orthogonality

$$\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0 \qquad \text{in } (0, T) \times \Omega. \qquad (1.5)$$

The immediate consequence $\partial_t(|\boldsymbol{m}|^2) = 2\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$ shows that the evolution dynamics is constrained, as the unit-length $|\boldsymbol{m}^0| \equiv 0$ in Ω is intrinsically preserved by $(\mathcal{LL})$, i.e., the magnetization always takes values on the unit sphere

$$|\boldsymbol{m}(t, \boldsymbol{x})| = 1 \qquad \text{for all } (t, \boldsymbol{x}) \in (0, T) \times \Omega. \qquad (1.6)$$

Geometrically speaking, the two terms on the right-hand side of $(\mathcal{LL})$ describe the combination of two effects: Considering a fixed point $\boldsymbol{x} \in \Omega$, the first term induces a precession $t \mapsto \boldsymbol{m}(t, \boldsymbol{x})$ around the effective field $t \mapsto \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}(t, \cdot))(\boldsymbol{x})$, while the second term describes a damping $t \mapsto \boldsymbol{m}(t, \boldsymbol{x})$ towards $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}(t, \cdot))(\boldsymbol{x})$. In combination we obtain a damped precession dynamics of $\boldsymbol{m}$ around and towards $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})$ resulting in a relative spiral-like motion until they align in the limit. This description is illustrated in Figure 1.1 for a coordinate system aligned with the effective field. We note that this illustration is only schematic, as

(a) Precession   (b) Damping   (c) Damped precession
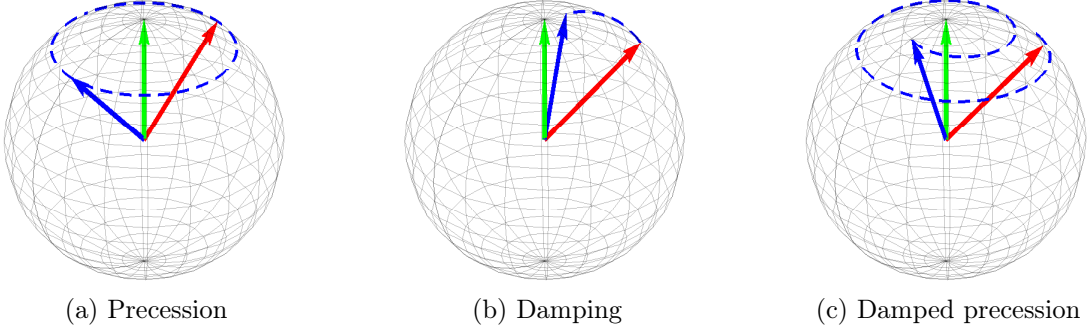
Figure 1.1: The blue dashed line illustrates the magnetization dynamics $t \mapsto \boldsymbol{m}(t, \boldsymbol{x})$ for some fixed $\boldsymbol{x} \in \Omega$ relative to $t \mapsto \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}(t, \cdot))(\boldsymbol{x})$ in green. Starting from $\boldsymbol{m}^0(\boldsymbol{x})$ in red, after some time $\tau > 0$, the magnetization obtains the blue value $\boldsymbol{m}(\tau, \boldsymbol{x})$.

the effective field $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m})$ nonlocally depends on the magnetization $\boldsymbol{m}$ and thus evolves in time.

Next, we observe that taking the scalar product of $(\mathcal{LL})$ with $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m})$ followed by integration over $(0, T) \times \Omega$ and using (1.2), we obtain the continuous energy law

$$\widehat{\mathcal{E}}(\boldsymbol{m}(T)) + \frac{\alpha}{1 + \alpha^2} \int_0^T \|\boldsymbol{m}(t) \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}(t))\|_{\boldsymbol{L}^2(\Omega)}^2 \, \mathrm{d}t = \widehat{\mathcal{E}}(\boldsymbol{m}^0) \,, \tag{1.7}$$

revealing the dissipative nature of LLG, where $\widehat{\mathcal{E}} := \mathcal{E}/(\mu_0 M_s^2)$ denotes the rescaled energy. We note that in the presence of a non-constant in-time external applied field $\boldsymbol{f} : (0, T) \times \Omega \to \mathbb{R}^3$, the left-hand side of (1.7) is augmented by the term $\int_0^T \langle \partial_t \boldsymbol{f}(t), \boldsymbol{m}(t) \rangle \, \mathrm{d}t$; see (1.13) below.

As we are dealing with a highly nonlinear partial differential equation, theory on strong solutions of LLG is rather limited: Under severe assumptions on the initial condition $\boldsymbol{m}^0$, existence of a local in-time strong solution was established in [CF01a]. Global in-time existence of strong solutions has only been shown in 2D [CF01a], or if $\Omega = \mathbb{R}^3$ [CF01b] (i.e., without boundary conditions), both, under severe smallness assumptions on the initial data $\boldsymbol{m}^0$. In the work [FT17b], the authors proved that local strong solutions are arbitrarily smooth, if the initial data $\boldsymbol{m}^0$ is sufficiently smooth and sufficiently close to a constant function. In the seminal work on weak solutions for LLG [AS92], it was shown that weak solutions exist globally in time, but are non-unique in general — more precisely, there even exist initial data $\boldsymbol{m}^0 \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ allowing for an uncountably infinite number of weak solutions. Weak-strong uniqueness (i.e., if a strong solution exists up to time T, then any weak solution coincides with it until time T) was established for the exchange-only case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \boldsymbol{\Delta} \boldsymbol{m}$ and $\Omega = \mathbb{R}^3$ in [DS14], and extended to $\Omega \subset \mathbb{R}^3$ and more general effective fields in [DIP20]. In particular, no analytical means to explicitly integrate LLG in time $t \mapsto \boldsymbol{m}(t)$ is available in the literature. Hence, to study micromagnetic processes, besides possibly costly real-life experiments, one can resort to numerical integration of LLG. This gives rise to the challenging field of computational micromagnetics, aiming to approximate the dynamics of LLG by fully discrete numerical schemes.

Throughout the thesis, 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 $\langle \cdot, \cdot \rangle$ the scalar product in $\boldsymbol{L}^2(\Omega)$. For a matrix-valued function $\boldsymbol{G} \colon \Omega \to \mathbb{R}^{3\times3}$ and a vector-valued function $\boldsymbol{g} \colon \Omega \to \mathbb{R}^3$, the cross product $\boldsymbol{G} \times \boldsymbol{g} \colon \Omega \to \mathbb{R}^{3\times3}$ is to be understood column-wise. Further, we use upper-case letters for physical quantities, e.g., $\boldsymbol{M}$ and $\boldsymbol{H}_{\text{eff}}$, and lower-case letters for their rescaled, nondimensional counterparts, e.g., $\boldsymbol{m}$ and $\boldsymbol{h}_{\text{eff}}$.

1.2 The micromagnetic energy

In dynamical micromagnetism, the magnetization evolves towards an equilibrium state, a configuration (locally) minimizing the system's total energy $\mathcal{E}(\boldsymbol{m})$. The total energy, or *Gibbs free energy* $\mathcal{E}(\boldsymbol{m})$, comprises several energy contributions.

Heisenberg exchange energy

We start with the symmetric exchange interaction, which we already discussed briefly in Section 1.1. The tendency of neighbouring magnetic moments to align locally with each other is described by the symmetric exchange energy contribution

$$\mathcal{E}_{\text{ex}}(\boldsymbol{m}) = A \int_\Omega |\boldsymbol{\nabla}\boldsymbol{m}|^2 \mathrm{d}\boldsymbol{x} \,, \tag{1.8a}$$

which is minimized for constant vector fields $\boldsymbol{m}$. The exchange energy yields the highest-order effective field contribution

$$-\frac{1}{\mu_0 M_{\text{s}}^2} \frac{\delta\mathcal{E}_{\text{ex}}(\boldsymbol{m})}{\delta\boldsymbol{m}} = \ell_{\text{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m} \,, \tag{1.8b}$$

which makes it the fundamental contribution in LLG analysis. In fact, in many of the proofs in the analysis of LLG, as well as in the analysis of numerical schemes for LLG, the central argument is to control the exchange energy of the micromagnetic system. Hence, results are often first formulated for the exchange-only setting $\mathcal{E}(\boldsymbol{m}) = \mathcal{E}_{\text{ex}}(\boldsymbol{m})$, and only later extended to more general energy contributions.

As the Laplacian is a local operator, from a computational point of view its discretization and evaluation are straightforward and do not pose any major issues. However, as it is the highest-order contribution, like for the heat equation, a stable numerical scheme must employ implicit in-time integration on it.

Anisotropy energy

The crystalline structure of a ferromagnetic body can favor one or more directions, called *easy axes*, where the magnetization prefers to align with in order to minimize energy. In case of a uniaxial anisotropy with easy axis $\boldsymbol{a} \in \mathbb{S}^2$, it costs energy to move the magnetization from this easy axis, which is described by the uniaxial anisotropy contribution

$$\mathcal{E}_{\text{ani}}(\boldsymbol{m}) = K \int_\Omega \left[1 - (\boldsymbol{a} \cdot \boldsymbol{m})^2\right] \mathrm{d}\boldsymbol{x}, \tag{1.9a}$$

where $K > 0$ denotes the material-dependent anisotropy constant (in J/m^3). Uniaxial anisotropy leads to the effective field contribution

$$\boldsymbol{h}_{\mathrm{ani}}(\boldsymbol{m}) := -\frac{1}{\mu_0 M_{\mathrm{s}}^2} \frac{\delta \mathcal{E}_{\mathrm{ani}}(\boldsymbol{m})}{\delta \boldsymbol{m}} = \frac{2K}{\mu_0 M_{\mathrm{s}}^2} (\boldsymbol{a} \cdot \boldsymbol{m}) \boldsymbol{a} \,, \tag{1.9b}$$

which is a zero-order, local, linear, and self-adjoint operator. For these reasons it does not pose any analytical or computational difficulties. In our presentation it is covered, among other terms, by a bounded, linear, and self-adjoint operator $\boldsymbol{\pi} \colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$.

Magnetostatic energy

According to the Maxwell equations, the magnetization induces a magnetic field. As the electromagnetic wavelength is much larger than the dimensions of a ferromagnet, this so-called *demagnetizing field* or *stray field* $\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m})$ is appropriately described by the magneto-static Maxwell equations

$$\nabla \cdot \boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m}) = -M_{\mathrm{s}} \nabla \cdot (\chi_\Omega \boldsymbol{m}) \qquad \text{and} \qquad \nabla \times \boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m}) = 0 \qquad \text{in } \mathbb{R}^3,$$

with χ_Ω the characteristic function on Ω. The stray field can be written as $\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m}) = -\nabla u|_\Omega$, where the magnetostatic potential $u \colon \mathbb{R}^3 \to \mathbb{R}$ solves the full-space transmission problem

$$-\Delta u = -M_{\mathrm{s}} \nabla \cdot \boldsymbol{m} \quad \text{in } \Omega, \tag{1.10a}$$

$$-\Delta u = 0 \quad \text{in } \mathbb{R}^3 \setminus \overline{\Omega}, \tag{1.10b}$$

$$u^{\mathrm{ext}} - u^{\mathrm{int}} = 0 \quad \text{on } \partial\Omega, \tag{1.10c}$$

$$(\nabla u^{\mathrm{ext}} - \nabla u^{\mathrm{int}}) \cdot \boldsymbol{n} = -M_{\mathrm{s}} \boldsymbol{m} \cdot \boldsymbol{n} \quad \text{on } \partial\Omega, \tag{1.10d}$$

$$u(\boldsymbol{x}) = \mathcal{O}(1/|\boldsymbol{x}|) \quad \text{as } |\boldsymbol{x}| \to \infty, \tag{1.10e}$$

where u^{ext} and u^{int}, respectively, denote the trace of $u|_{\mathbb{R}^3 \setminus \overline{\Omega}}$ and $u|_\Omega$ on $\partial\Omega$. The operator $\boldsymbol{H}_{\mathrm{s}}$ is linear, self-adjoint and yields the energy of the magnetization interacting with its own demagnetizing field

$$\mathcal{E}_{\mathrm{stray}}(\boldsymbol{m}) = \frac{\mu_0}{2} \int_{\mathbb{R}^3} |\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m})|^2 \, \mathrm{d}\boldsymbol{x} = -\frac{\mu_0 M_{\mathrm{s}}}{2} \int_\Omega \boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m}) \cdot \boldsymbol{m} \, \mathrm{d}\boldsymbol{x} \,. \tag{1.11a}$$

This leads to the effective field contribution

$$\boldsymbol{h}_{\mathrm{s}}(\boldsymbol{m}) = -\frac{1}{\mu_0 M_{\mathrm{s}}^2} \frac{\delta \mathcal{E}_{\mathrm{stray}}(\boldsymbol{m})}{\delta \boldsymbol{m}} = \frac{\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m})}{M_{\mathrm{s}}} \,, \tag{1.11b}$$

which is a nonlocal, linear, and self-adjoint operator [Pra04]. Consequently, it does not pose any analytical difficulties and in our presentation, like the anisotropy contribution, it is covered by the bounded, linear, and self-adjoint operator $\boldsymbol{\pi} \colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$.

In practice, however, the nonlocality of $\boldsymbol{h}_{\mathrm{s}}$ poses a serious computational challenge. One possibility to reliably approximate $\boldsymbol{h}_{\mathrm{s}}(\boldsymbol{m})$, which we also employ in Chapter 2–6, is proposed in [FK90], where a hybrid finite element / boundary element method (FEM/BEM) is formulated. There, the potential is decomposed as $u = u_1 + u_2$ to approximate u_1 by the

FEM solution to a Neumann problem, while u_2 is approximated by the FEM solution to an inhomogeneous Dirichlet problem. The expensive part is the computation of the boundary data for the inhomogeneous Dirichlet problem, which is obtained from a discretization of the double-layer potential evaluated on u_1 restricted to $\partial\Omega$. We note that the approach from [FK90] leads to a discrete operator $\boldsymbol{\pi}_h\colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ introducing an additional consistency error $(\boldsymbol{\pi} - \boldsymbol{\pi}_h)(\boldsymbol{m})$. Alternatively, one could employ a coupled FEM-BEM method with the drawback of solving the more expensive coupled system.

Zeeman energy

Under influence of an externally applied magnetic field $\boldsymbol{H}_{\mathrm{ext}}$ (in A/m), the magnetization tends to align in the direction of the applied field. This tendency is described by the so-called *Zeeman energy contribution*

$$\mathcal{E}_{\mathrm{ext}}(\boldsymbol{m}) = -\mu_0 M_{\mathrm{s}} \int_\Omega \boldsymbol{H}_{\mathrm{ext}} \cdot \boldsymbol{m} \, \mathrm{d}\boldsymbol{x} \,, \tag{1.12a}$$

which leads to the effective field contribution

$$\boldsymbol{f} = -\frac{1}{\mu_0 M_{\mathrm{s}}^2} \frac{\delta \mathcal{E}_{\mathrm{ext}}(\boldsymbol{m})}{\delta \boldsymbol{m}} = \frac{\boldsymbol{H}_{\mathrm{ext}}}{M_{\mathrm{s}}} \,, \tag{1.12b}$$

which is independent of the magnetization $\boldsymbol{m}$ and bounded, as usually $\boldsymbol{H}_{\mathrm{ext}}$ is assumed to be bounded in $\boldsymbol{L}^2(\Omega)$. Hence, it does not lead to analytical or computational difficulties in our presentation. However, although in theory no additional difficulty is introduced by the applied field, it yields some technical pitfalls in the presentation: In presence of an applied field, the effective field is not linear anymore, i.e., with our notation there holds

$$\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{u}) + \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{v}) = \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{f} \,,$$

which has to be kept in mind. Moreover, if $\boldsymbol{H}_{\mathrm{ext}}$ is not constant in time, additional energy is supplied to or drawn from the system

$$-\frac{1}{\mu_0 M_{\mathrm{s}}^2} \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(\boldsymbol{m}, \boldsymbol{f}) = \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) \cdot \partial_t \boldsymbol{m} + \boldsymbol{m} \cdot \partial_t \boldsymbol{f} \,, \tag{1.13}$$

which manifests in the energy law (1.7), i.e., the term $\int_0^T \langle \partial_t \boldsymbol{f}(t), \boldsymbol{m}(t) \rangle \, \mathrm{d}t$ has to be added to the left-hand side of (1.7).

Dzyaloshinskii–Moriya interaction energy

The antisymmetric exchange interaction yields the most important energy contribution for the enucleation and the stabilization of chiral magnetic skyrmions. It was phenomenologically introduced to the LLG model by Dzyaloshinskii [Dzy58] and Moriya [Mor60], and comprises a linear combination of the so-called *Lifshitz invariants*, the components of the chirality tensor $\boldsymbol{m} \times \boldsymbol{\nabla}\boldsymbol{m}\colon \Omega \to \mathbb{R}^{3\times3}$. Taking the trace of the chirality tensor, we obtain the so-called *bulk Dzyaloshinskii–Moriya (DMI) energy contribution*

$$\mathcal{E}_{\mathrm{dm}}(\boldsymbol{m}) = D \int_\Omega \left(\nabla \times \boldsymbol{m} \right) \cdot \boldsymbol{m} \, \mathrm{d}\boldsymbol{x} \,, \tag{1.14a}$$

where $D \in \mathbb{R}$ is the DMI constant (in J/m^2). With the DMI length $\ell_{\mathrm{dm}} = 2D/(\mu_0 M_{\mathrm{s}}^2)$ (measured in m), the bulk DMI energy leads to the effective field contribution

$$\boldsymbol{h}_{\mathrm{dm}}(\boldsymbol{m}) = -\frac{1}{\mu_0 M_{\mathrm{s}}^2} \frac{\delta \mathcal{E}_{\mathrm{dm}}(\boldsymbol{m})}{\delta \boldsymbol{m}} = -\ell_{\mathrm{dm}} \nabla \times \boldsymbol{m} \,, \tag{1.14b}$$

which is a first-order, local, linear, non-self-adjoint operator. As $\boldsymbol{h}_{\mathrm{dm}}(\boldsymbol{m})$ includes spatial derivatives, in presence of DMI the imposed natural boundary conditions of the LLG system (1.4) are adjusted in order to be consistent with the Euler–Lagrange equations obtained from the energy minimization problem

$$\min_{|\boldsymbol{m}|=1} \mathcal{E}(\boldsymbol{m}) \,.$$

In case of the bulk DMI energy contribution, this leads to the consistent boundary conditions

$$\ell_{\mathrm{ex}}^2 \partial_{\boldsymbol{n}} \boldsymbol{m} = -\frac{\ell_{\mathrm{dm}}}{2} \boldsymbol{m} \times \boldsymbol{n} \qquad \text{on } (0, T) \times \partial\Omega \,, \tag{1.15}$$

replacing the homogeneous Neumann boundary conditions (1.4).

Different to the symmetric exchange energy $\mathcal{E}_{\mathrm{ex}}(\boldsymbol{m})$, the antisymmetric exchange energy $\mathcal{E}_{\mathrm{dm}}(\boldsymbol{m})$ can be positive or negative. Despite the lower-order of the contribution, this lack of a sign complicates the analysis, e.g., when deriving $\boldsymbol{H}^1(\Omega)$-bounds on $\boldsymbol{m}$ from boundedness of the total energy. Due to its locality, the DMI contribution does not pose major computational issues. However, in a variational formulation of LLG boundary integrals obtained from (1.15) must be reformulated as volume integrals by Green's formula; see, e.g., (2.12) in Section 2.

DMI is essential for an appropriate description of magnetic skyrmions in ferromagnetic bodies. These vortex-like structures are topologically protected and occur as (local) minimizers of the micromagnetic energy. Lately, the investigation of skyrmion dynamics has attained quite some attention, as in [SCR$^+$13], by numerical investigations the authors showed that there are nanostructures allowing for two topologically different stable states — a skyrmion state and a uniform state — both locally minimizing the total micromagnetic energy; see Figure 1.2. This bi-stability in presence of DMI lead to great interest in the simulation of skyrmion dynamics, e.g., in the research for future memory technologies, where the two possible states of a logical bit could be encoded as the presence or absence of a magnetic skyrmion.

For the remainder of this chapter, we restrict our presentation to the classical energy contributions

$$\mathcal{E}(\boldsymbol{m}) = \mathcal{E}_{\mathrm{ex}}(\boldsymbol{m}) + \mathcal{E}_{\mathrm{ani}}(\boldsymbol{m}) + \mathcal{E}_{\mathrm{stray}}(\boldsymbol{m}) + \mathcal{E}_{\mathrm{ext}}(\boldsymbol{m}) \,.$$

Additionally, from now on we consider the rescaled energy $\widehat{\mathcal{E}} := \mathcal{E}/(\mu_0 M_{\mathrm{s}}^2)$, but to simplify the notation it will be denoted by $\mathcal{E}$; see, e.g., (1.18) and (1.20) below. Collecting the anistropy contribution and the stray field contribution in the bounded, linear, and self-adjoint operator $\boldsymbol{\pi} \colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$, the classical energy yields the effective field

$$\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta} \boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f} \,.$$

For the extension of the numerical integrators in presence of DMI, we refer to Chapter 2 and Chapter 3 of this thesis.
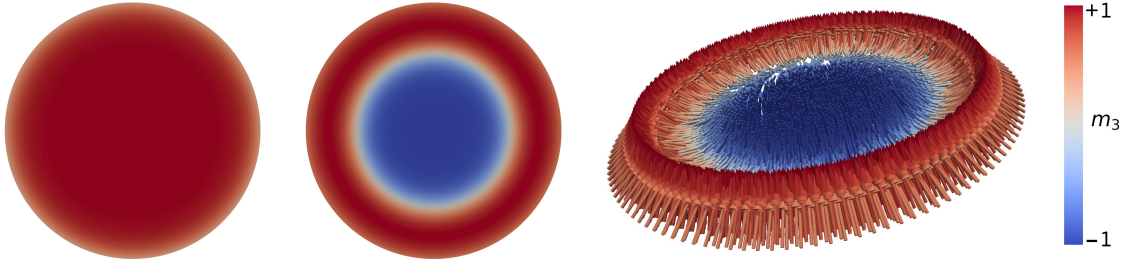
Figure 1.2: Two topologically different stable states in a ferromagnetic nanodisk inspiring technological applications. Left: Almost uniform equilibrium state with slight inward tilt on the boundary of the disk, viewed from top. Middle: Skyrmion equilibrium state, viewed from top. Right: Skyrmion equilibrium state, 3D view diagonally from above. For details, see the numerical studies in Section 2.4.2 and Section 3.4.1.

1.3 Landau–Lifshitz–Gilbert equation

The Landau–Lifshitz–Gilbert equation is a phenomenological nonlinear parabolic partial differential equation, which is a well-accepted model for the description of micromagnetic phenomena. We briefly touch on the historical side of the model, where we follow the presentation in the review article [Cim08].

In 1935, Lev D. Landau and Evgeny M. Lifshitz proposed the first dynamical model to describe precessional motion of the magnetization including energy dissipation [LL35]. With the physical vector fields $\boldsymbol{M}$ and $\boldsymbol{H}_{\text{eff}}$ (both in A/m), physical constants and parameters γ_0 denoting the rescaled gyromagnetic ratio (in m/(A s)), the dimensionless damping parameter α, and the saturation magnetization M_{s} (in A/m), the proposed system reads

$$\partial_t \boldsymbol{M} = -\gamma_0\, \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}} - \gamma_0 \frac{\alpha}{M_{\text{s}}}\, \boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}_{\text{eff}})\,.$$

It includes a conservative rotational precession term matching the theory of gyromagnetic precession, and a phenomenological term describing energy dissipation by forcing the magnetization towards the effective field. Mathematically, it is more convenient to work with normalized quantities, i.e., with $\boldsymbol{m} := \boldsymbol{M}/M_{\text{s}}$ and the physical parameters $\alpha_L := \gamma_0 \alpha$ and $\beta_L := \gamma_0$ we obtain

$$\partial_t \boldsymbol{m} = -\beta_L\, \boldsymbol{m} \times \boldsymbol{H}_{\text{eff}} - \alpha_L\, \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{H}_{\text{eff}})\,. \tag{1.16}$$

Later in 1955, Thomas L. Gilbert proposed to augment the conservative precession by viscous damping effects consistently derived from the Rayleigh dissipation function [Gil55]. This led him to the equation

$$\partial_t \boldsymbol{m} = -\beta_G\, \boldsymbol{m} \times \boldsymbol{H}_{\text{eff}} + \alpha_G\, \boldsymbol{m} \times \partial_t \boldsymbol{m}\,, \tag{1.17}$$

incorporating the physical parameters α_G and β_G. In this model, the damped precession dynamics is incorporated implicitly, i.e., in contrast to (1.16) the two terms on the right-hand side are not explicitly orthogonal to each other by the cross product structure. Nevertheless, for $\beta_G \neq 0$ the two equations are mathematically equivalent with $\alpha_G = \alpha_L/\beta_L$ and $\beta_G = (\alpha_L^2 + \beta_L^2)/\beta_L$ as shown in [Cim08].

1.3.1 Equivalent formulations

In (1.17), nondimensionalization of the time via the transformation $t \rightsquigarrow (1+\alpha^2)M_{\mathrm{s}}\gamma_0\, t$ yields the dimensionless Gilbert form $(\mathcal{G})$ below, where $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) := \boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m})/M_{\mathrm{s}}$. The numerical schemes investigated in this thesis are based on the following three different dimensionless forms of LLG.

- The *Landau–Lifshitz form* of LLG reads

$$(1 + \alpha^2)\partial_t\boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) - \alpha\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})), \qquad (\mathcal{LL})$$

  and gives rise to the family of recent predictor-corrector methods from [KW18] presented in Section 1.5.4, Section 1.6.4, and investigated in Chapter 5.

- The *Gilbert form* of LLG reads

$$\partial_t\boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) + \alpha\,\boldsymbol{m} \times \partial_t\boldsymbol{m}, \qquad (\mathcal{G})$$

  and gives rise to the family of midpoint schemes [BP06] presented in Section 1.5.2, Section 1.6.2, and investigated in Chapter 3.

- The so-called *alternative form* of LLG reads

$$\alpha\,\partial_t\boldsymbol{m} + \boldsymbol{m} \times \partial_t\boldsymbol{m} = \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) - (\boldsymbol{m} \cdot \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}))\boldsymbol{m}, \qquad (\mathcal{A})$$

  and gives rise to the family of tangent plane schemes [Alo08] presented in Section 1.5.3, Section 1.6.3, and investigated in Chapter 2 and Chapter 4.

The Landau–Lifshitz form $(\mathcal{LL})$ and the Gilbert form $(\mathcal{G})$ are equivalent to each other, while the alternative form $(\mathcal{A})$ does not explicitly preserve unit-length (1.6), but is equivalent to the other forms if additionally $|\boldsymbol{m}| = 1$ is assumed. For the Gilbert form $(\mathcal{G})$, the orthogonality (1.5) and intrinsic preservation of unit-length (1.6) are obtained analogously to the computations in Section 1.1 for the Landau–Lifshitz form $(\mathcal{LL})$.

Taking the scalar product of $(\alpha d_t\boldsymbol{m} + \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}))$ with $(\mathcal{G})$ gives a slightly different (but equivalent) form of the energy law (1.7), namely

$$\mathcal{E}(\boldsymbol{m}(T)) + \alpha\int_0^T \|\partial_t\boldsymbol{m}(t)\|_{\boldsymbol{L}^2(\Omega)}^2\,\mathrm{d}t = \mathcal{E}(\boldsymbol{m}^0), \qquad (1.18)$$

where we again assumed an external applied field, which is constant in time; see (1.13). If orthogonality (1.5) is assumed, the same energy law is obtained from the alternative form $(\mathcal{A})$ by taking the scalar product with $\alpha\partial_t\boldsymbol{m}$. We briefly summarize how the three forms are derived from each other using the triple product expansion formula

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c} \qquad \text{for all } \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3.$$

- $(\mathcal{G})$ is obtained from the Landau–Lifshitz form via $[(\mathcal{LL}) - \alpha\boldsymbol{m} \times (\mathcal{LL})]/(1+\alpha^2)$.

- $(\mathcal{LL})$ is obtained from the Gilbert form via $[(\mathcal{G}) + \alpha\boldsymbol{m} \times (\mathcal{G})]$.

- $(\mathcal{A})$ is obtained from the Gilbert form via $[\boldsymbol{m} \times (\mathcal{G})]$.

- $(\mathcal{LL})$ and $(\mathcal{G})$ can be derived from the alternative form, if additionally $|\boldsymbol{m}| = 1$ (or, equivalently $\boldsymbol{m} \cdot \partial_t\boldsymbol{m} = 0$) is assumed. Then, $(\mathcal{G})$ is obtained via $[\boldsymbol{m} \times (\mathcal{A})]$.

1.3.2 Weak solutions to LLG

We give the notion of a weak solution of LLG, which extends the one introduced in [AS92]. We refer to Chapter 2 and Chapter 3 for the notion of a weak solution in presence of DMI.

Definition 1.3.1. *Let $\boldsymbol{m}^0 \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ and $\boldsymbol{f} \in C^1([0,T]; \boldsymbol{L}^2(\Omega))$. A vector field $\boldsymbol{m} \colon (0,T) \times \Omega \to \mathbb{R}^3$ is called a (local) weak solution of $(\mathcal{LL})$ with (1.3)–(1.4), if the following properties are satisfied:*

(i) $\boldsymbol{m} \in \boldsymbol{H}^1((0,T) \times \Omega) \cap L^\infty(0,T; \boldsymbol{H}^1(\Omega))$ with $|\boldsymbol{m}| = 1$ a.e. in $(0,T) \times \Omega$;

(ii) $\boldsymbol{m}(0) = \boldsymbol{m}^0$ in the sense of traces;

(iii) for all $\boldsymbol{w} \in \boldsymbol{H}^1((0,T) \times \Omega)$, it holds that

$$
\int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{w}(t) \rangle \, \mathrm{d}t - \alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{w}(t) \rangle \, \mathrm{d}t
$$
$$
= \ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{\nabla} \boldsymbol{m}(t), \boldsymbol{\nabla} \boldsymbol{w}(t) \rangle \, \mathrm{d}t - \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{\pi}(\boldsymbol{m}(t)), \boldsymbol{w}(t) \rangle \, \mathrm{d}t \quad (1.19)
$$
$$
- \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{f}(t), \boldsymbol{w}(t) \rangle \, \mathrm{d}t;
$$

(iv) it holds that

$$
\mathcal{E}(\boldsymbol{m}(T)) + \alpha \int_0^T \|\partial_t \boldsymbol{m}(t)\|^2 \mathrm{d}t + \int_0^T \langle \partial_t \boldsymbol{f}(t), \boldsymbol{m}(t) \rangle \, \mathrm{d}t \leq \mathcal{E}(\boldsymbol{m}^0). \quad (1.20)
$$

We note that (1.19) is a variational formulation in space-time of the Gilbert form $(\mathcal{G})$ of LLG, and that (1.20) is a weaker version of the energy law (1.18). In the literature, a vector field $\boldsymbol{m} \in \boldsymbol{L}^\infty((0,\infty); \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ is called a global weak solution of $(\mathcal{LL})$ with (1.3)–(1.4), if it is a local weak solution in the sense of Definition 1.3.1 for any $T > 0$.

1.4 Discretization

Uniform time discretization

Given $J \in \mathbb{N}$, we consider the uniform partition $\{t_i = iT/J \colon i = 0, \ldots, J\}$ of the time interval $[0,T]$ with time-step size $k := T/J > 0$, i.e., $t_i = ik$ for all $i = 0, \ldots, J$. Given a finite sequence of corresponding spatial functions $\{\boldsymbol{u}^i\}_{i=0}^J$, we define

$$
\boldsymbol{u}^{i+1/2} := \frac{\boldsymbol{u}^{i+1} + \boldsymbol{u}^i}{2} \quad \text{and} \quad d_t \boldsymbol{u}^{i+1} := \frac{\boldsymbol{u}^{i+1} - \boldsymbol{u}^i}{k} \quad \text{for all } i = 0, \ldots J - 1.
$$

Space discretization by finite elements

In this thesis, the focus is on finite element-based numerical schemes for LLG. Compared to finite difference-based schemes, which are computationally cheaper in general, finite

elements are much more flexible in the discretization and allow for computations of far more complicated domains $\Omega \subset \mathbb{R}^3$.

For the spatial discretization, we consider a regular tetrahedral triangulation $\mathcal{T}_h$ of Ω with mesh size $h > 0$. We denote by $\mathcal{N}_h$ the set of vertices of $\mathcal{T}_h$ and by $\{\phi_{\boldsymbol{z}}\}_{\boldsymbol{z} \in \mathcal{N}_h}$ the classical nodal basis of the space $\mathcal{S}^1(\mathcal{T}_h)$ of $\mathcal{T}_h$-piecewise linear and globally continuous discrete functions, i.e., $\phi_{\boldsymbol{z}}(\boldsymbol{z}') = \delta_{\boldsymbol{z},\boldsymbol{z}'}$ for all $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{N}_h$. With $\{\boldsymbol{e}_j\}_{j=1,2,3}$ the standard basis of $\mathbb{R}^3$, $\{\phi_{\boldsymbol{z}} \boldsymbol{e}_j\}_{\boldsymbol{z} \in \mathcal{N}_h, j=1,2,3}$ gives a basis of $\mathcal{S}^1(\mathcal{T}_h)^3$. Note that $\mathcal{S}^1(\mathcal{T}_h)^3$ is a $3N$-dimensional space, with N denoting the number of vertices in $\mathcal{N}_h$. Inspired by (1.6) and (1.5), we introduce the *set of admissible discrete magnetizations*

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

and, for $\boldsymbol{m}_h \in \boldsymbol{\mathcal{M}}_h$, the *discrete tangent space*

$$\boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h] := \left\{ \boldsymbol{\varphi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3 \colon \boldsymbol{m}_h(\boldsymbol{z}) \cdot \boldsymbol{\varphi}_h(\boldsymbol{z}) = 0 \text{ for all } \boldsymbol{z} \in \mathcal{N}_h \right\}. \tag{1.21}$$

We consider the nodal interpolant $\mathcal{I}_h \colon C^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)$, which is defined by $\mathcal{I}_h(v) = \sum_{\boldsymbol{z} \in \mathcal{N}_h} v(\boldsymbol{z}) \phi_{\boldsymbol{z}}$ for all $v \in C^0(\overline{\Omega})$. We denote the vector-valued realization of the nodal interpolant by $\boldsymbol{\mathcal{I}}_h \colon \boldsymbol{C}^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)^3$. On $\boldsymbol{C}^0(\overline{\Omega})$, besides the standard $\boldsymbol{L}^2(\Omega)$-scalar product $\langle \cdot, \cdot \rangle$, we consider the *mass-lumped* scalar product $\langle \cdot, \cdot \rangle_h$ defined by

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

Using the definition of the nodal interpolant, we see that

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle_h = \sum_{\boldsymbol{z} \in \mathcal{N}_h} \beta_{\boldsymbol{z}} \, \boldsymbol{u}(\boldsymbol{z}) \cdot \boldsymbol{w}(\boldsymbol{z}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{w} \in \boldsymbol{C}^0(\overline{\Omega}), \tag{1.22}$$

where $\beta_{\boldsymbol{z}} := \int_\Omega \phi_{\boldsymbol{z}} \, \mathrm{d}\boldsymbol{x} > 0$ for all $\boldsymbol{z} \in \mathcal{N}_h$. For discrete functions, the induced norm $\|\cdot\|_h$ is equivalent to the standard $\boldsymbol{L}^2(\Omega)$-norm; see [Bar15, Lemma 3.9], i.e., it holds that

$$\|\boldsymbol{w}_h\|_{\boldsymbol{L}^2(\Omega)} \leq \|\boldsymbol{w}_h\|_h \leq \sqrt{5} \|\boldsymbol{w}_h\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{1.23}$$

We define the (negative) *discrete Laplacian* $-\boldsymbol{\Delta}_h \colon \boldsymbol{H}^1(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)^3$ by

$$-\langle \boldsymbol{\Delta}_h \boldsymbol{w}, \boldsymbol{w}_h \rangle_h = \langle \boldsymbol{\nabla} \boldsymbol{w}, \boldsymbol{\nabla} \boldsymbol{w}_h \rangle \quad \text{for all } \boldsymbol{w} \in \boldsymbol{H}^1(\Omega) \text{ and } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{1.24}$$

Some integrators enforce the nodewise unit-length constraint via a nodewise projection onto the unit sphere $\mathbb{S}^2$. The stability analysis of those integrators relies on the so-called *angle condition*: It is satisfied, if all off-diagonal entries of the stiffness matrix $(a_{\boldsymbol{z},\boldsymbol{z}'})_{\boldsymbol{z},\boldsymbol{z}' \in \mathcal{N}_h}$ are nonpositive, i.e., if it holds that

$$a_{\boldsymbol{z},\boldsymbol{z}'} := \langle \nabla \boldsymbol{w}_{\boldsymbol{z}'}, \nabla \boldsymbol{w}_{\boldsymbol{z}} \rangle \leq 0 \quad \text{for all } \boldsymbol{z}, \boldsymbol{z}' \in \mathcal{N}_h \text{ with } \boldsymbol{z} \neq \boldsymbol{z}'. \tag{1.25}$$

This requirement ensures that the nodal projection $\boldsymbol{w}_h \mapsto \boldsymbol{\mathcal{I}}_h \big[\boldsymbol{w}_h / |\boldsymbol{w}_h| \big]$ does not increase the exchange energy of a discrete function, i.e., it holds that

$$\big\| \boldsymbol{\nabla} \boldsymbol{\mathcal{I}}_h \big[\boldsymbol{w}_h / |\boldsymbol{w}_h| \big] \big\|_{\boldsymbol{L}^2(\Omega)} \leq \|\boldsymbol{\nabla} \boldsymbol{w}_h\|_{\boldsymbol{L}^2(\Omega)}, \tag{1.26}$$

for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$ with $|\boldsymbol{w}_h(\boldsymbol{z})| \geq 1$ for all $\boldsymbol{z} \in \mathcal{N}_h$; see [Bar05, Lemma 3.2]. Note that (1.26) is motivated by the pointwise estimate

$$\big|\boldsymbol{\nabla}\big[\boldsymbol{w}/|\boldsymbol{w}|\big]\big| \leq |\boldsymbol{\nabla}\boldsymbol{w}| \qquad \text{for all } \boldsymbol{w} \in \boldsymbol{C}^1(\overline{\Omega}) \text{ with } |\boldsymbol{w}| \geq 1 \text{ in } \Omega,$$

which follows from a straightforward computation. The assumption (1.25) is usually referred to as angle condition, because in 3D it is satisfied, e.g., if all dihedral angles of all tetrahedra of $\mathcal{T}_h$ are bounded above by $\pi/2$.

Weak sub-convergence

Under appropriate assumptions and given discretization parameters $h, k > 0$, the numerical schemes for LLG discussed in this thesis generate discrete approximations $\{\boldsymbol{m}_h^i\}_{i=0}^J$ at the uniform time-steps t_i. After an affine interpolation in time

$$\boldsymbol{m}_{hk}(t) := \frac{t - t_i}{k}\boldsymbol{m}_h^{i+1} + \frac{t_{i+1} - t}{k}\boldsymbol{m}_h^i \quad \text{for all } i = 0, \ldots, J - 1 \text{ and } t \in [t_i, t_{i+1}],$$

we obtain an approximation $\boldsymbol{m} \approx \boldsymbol{m}_{hk} \in \boldsymbol{H}^1((0,T)\times\Omega)$. Ideally, one would have $\boldsymbol{m}_{hk} \to \boldsymbol{m}$ strongly as $h, k \to 0$. For most of the numerical integrators discussed in this thesis, however, due to the strong nonlinearity of LLG, the approximations generated by the algorithm usually only satisfy a far weaker notion of convergence: Considering a family of approximations $\{\boldsymbol{m}_{hk}\}_{h,k>0} \subset \boldsymbol{H}^1((0,T) \times \Omega)$, weak convergence of a subsequence of $\{\boldsymbol{m}_{hk}\}_{h,k>0}$ towards a weak solution of LLG, in the sense of Definition 1.3.1, is established. In the presentation of the numerical integrators in Section 1.5 and Section 1.6 below, if not stated otherwise, we refer to this notion of convergence.

1.5 Three original finite element integrators

In the design of a numerical integrator, one aims to meet a number of analytical and practical requirements: First of all, analytically, a numerical scheme should be well-defined for any admissible input. Further, the discrete output should approximate the analytical solution to the continuous problem in some sense. Due to the nonlinearity of LLG, these first two fundamental demands already pose a non-trivial task. Additionally for LLG, one desires to resemble the intrinsic characteristics of the continuous magnetization dynamics on a discrete level, i.e., the unit-length constraint (1.6) and the energy law (1.7).

Let us elaborate on how these fundamental analytical challenges are addressed by the three different (families of) numerical schemes discussed in this thesis. Later, in Section 1.6, further requirements towards practical feasibility of the schemes are formulated, leading to extensions of the three original approaches.

For some $J \in \mathbb{N}$ the magnetization is approximated at uniformly distributed time-steps $\{t_i := iT/J : i = 0, \ldots, J\} \subset [0, T]$. The continuous formulation is reduced to a finite dimensional discrete problem by considering lowest-order vector-valued $\boldsymbol{H}^1(\Omega)$-conforming finite elements $\mathcal{S}^1(\mathcal{T}_h)^3$ on a spatial triangulation $\mathcal{T}_h$ of the domain $\Omega \subset \mathbb{R}^3$ as introduced in

Section 1.4. Then, starting from a discretization $\boldsymbol{m}^0 \approx \boldsymbol{m}_h^0 \in \mathcal{S}^1(\mathcal{T}_h)^3$ of the initial state, one iterates through the discrete time steps

$$\boldsymbol{m}^0 \approx \boldsymbol{m}_h^0 \rightsquigarrow \boldsymbol{m}_h^1 \rightsquigarrow \cdots \rightsquigarrow \boldsymbol{m}_h^J \approx \boldsymbol{m}(T),$$

with iterates belonging to the discrete space $\boldsymbol{m}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$ for all $i = 0, \ldots, J$. The schemes discussed in this thesis now differ by the rule they prescribe to advance in time $\boldsymbol{m}_h^i \rightsquigarrow \boldsymbol{m}_h^{i+1}$.

1.5.1 Fundamental analytical requirements

The formulation of a fully discrete, practically applicable numerical scheme for LLG poses several difficulties. Above, we identified three fundamental challenges in the design of a numerical scheme for LLG: Dealing with the nonlinearity of LLG; respecting the unit-length constraint (1.6); and reflecting the energy law (1.7).

First of all, the nonlinearity of LLG has to be handled appropriately as on the computational level we are essentially limited to the solution of linear systems of equations of finite dimension. Due to the nonlinearity of LLG, the rigorous quantification of numerical approximation errors turns out to be non-trivial: While for all three original integrators presented in Section 1.5.2 through Section 1.5.4 below only weak convergence towards a weak solution for a subsequence of the computed approximates is known so far, experimentally the integrators are first- or second-order accurate in time, and first-order accurate (w.r.t. the energy error) in space.

Considering only those approximations, which satisfy the pointwise unit-length constraint $\boldsymbol{m}(\boldsymbol{x}) \in \mathbb{S}^2$ for all $\boldsymbol{x} \in \Omega$, would lead to a degeneration of the solution space, since in the finite element space $\mathcal{S}^1(\mathcal{T}_h)^3$ only a constant function can satisfy this side constraint. Hence, the continuous unit-length constraint is weakened to a nodewise constraint $\boldsymbol{m}(\boldsymbol{z}) \in \mathbb{S}^2$ for all vertices $\boldsymbol{z} \in \mathcal{N}_h$ of the triangulation $\mathcal{T}_h$.

Lastly, the fully discrete, practically applicable schemes recover only partially the continuous energy identity (1.7) on a discrete level either by a perturbed discrete energy equality, or by a discrete energy estimate. In both cases, eventually after imposing an appropriate coupling on the discretization parameters, the energy dissipation of the discrete scheme inherited from the continuous model allows to prove stability and convergence.

In this thesis three families of finite element-based schemes for the numerical integration of LLG are investigated:

- the midpoint scheme from [BP06],

- the tangent plane scheme from [Alo08],

- and two predictor-corrector methods recently proposed in [KW18].

In the following we (informally) describe the different approaches taken by the three (families of) numerical schemes discussed in this thesis to resolve the nonlinearity of LLG. Then, we elaborate how these approaches are combined with different update formulas $\boldsymbol{m}_h^i \rightsquigarrow \boldsymbol{m}_h^{i+1}$ preserving discrete unit-length. Finally, for each approach we discuss the effects on stability.

Nonlinearity. The three schemes use different approaches to handle the nonlinearity of LLG. For the midpoint scheme, the Gilbert form ($\mathcal{G}$) of LLG is discretized to arrive at a finite dimensional system of nonlinear equations. In practice, this nonlinear system is then linearized by an iterative solver.

On the other hand, the tangent plane scheme is based on the alternative form ($\mathcal{A}$) of LLG and in each time-step approximates the time derivative $\partial_t \boldsymbol{m}$. By using a variational formulation of ($\mathcal{A}$), which is posed in the tangent space of $\boldsymbol{m}$, the (highly nonlinear) second term on the right-hand side of ($\mathcal{A}$) is dropped. Hence, even with implicit integration in time, this variational form is linear in $\partial_t \boldsymbol{m}$ and the nonlinearity of LLG is circumvented as discretization of ($\mathcal{A}$) already leads to a discrete system of equations, which is linear in terms of the unknown approximating $\partial_t \boldsymbol{m}$.

The third family of integrators from [KW18] takes an approach in between: As for the tangent plane integrator, in each time-step the time derivative $\partial_t \boldsymbol{m}$ is approximated. But to compute this approximation, the Landau–Lifshitz form ($\mathcal{LL}$) is discretized. To arrive at a linear system of equations, only the effective field $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m})$ is integrated implicitly in time, while instances of the magnetization $\boldsymbol{m}$ are treated explicitly.

Unit-length constraint. The considered schemes take also different approaches for the preservation of the discrete unit-length constraint. For the tangent plane scheme nodewise unit-length is enforced explicitly by nodal projection. When advancing in time $\boldsymbol{m}_h^i \rightsquigarrow \boldsymbol{m}_h^{i+1}$, the tangent plane scheme uses linear time stepping to define an intermediate state $\hat{\boldsymbol{m}}_h^{i+1} := \boldsymbol{m}_h^i + k\boldsymbol{v}_h^i$, with $\boldsymbol{v}_h^i$ the computed approximation to the time derivative $\partial_t \boldsymbol{m}(t_i)$. Then, for each node $\boldsymbol{z} \in \mathcal{N}_h$, this intermediate state is renormalized by projection onto the unit sphere $\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \hat{\boldsymbol{m}}_h^{i+1}(\boldsymbol{z})/|\hat{\boldsymbol{m}}_h^{i+1}(\boldsymbol{z})| \in \mathbb{S}^2$.

Differently, the midpoint scheme guarantees preservation of nodewise unit-length implicitly as the discrete system defining the advancement $\boldsymbol{m}_h^i \rightsquigarrow \boldsymbol{m}_h^{i+1}$ ensures $|\boldsymbol{m}_h^{i+1}(\boldsymbol{z})| = |\boldsymbol{m}_h^i(\boldsymbol{z})|$ for all $\boldsymbol{z} \in \mathcal{N}_h$. The crucial ingredient for this identity to hold, is the mass-lumped integration rule (1.22) replacing exact spatial integration in the definition of the discrete (nonlinear) system of equations.

Again, for the family of integrators from [KW18] a mixed approach applies: Their first-order integrator, after computing $\boldsymbol{v}_h^i \approx \partial_t \boldsymbol{m}(t_i)$, applies the same projection update as the tangent plane scheme to enforce nodewise unit-length. Their second-order scheme, on the other hand, first computes $\boldsymbol{v}_h^i$ analogous to its first-order counterpart. Then, defining an intermediate state $\hat{\boldsymbol{m}}_h^{i+1/2} := \boldsymbol{m}_h^i + (k/2)\boldsymbol{v}_h^i$, another mass-lumped discrete variational formulation is solved to obtain $\boldsymbol{m}_h^{i+1}$, implicitly ensuring $|\boldsymbol{m}_h^{i+1}(\boldsymbol{z})| = |\boldsymbol{m}_h^i(\boldsymbol{z})|$. As for the midpoint scheme, this implicit preservation of discrete unit-length crucially relies on the mass-lumped integration rule (1.22).

Energy law and stability. A strength of the midpoint scheme in idealized form is its realization of a discrete energy equality resembling the continuous energy law (1.7). This is a consequence of the fully implicit approach taken to discretize ($\mathcal{G}$). Only for the practical midpoint schemes, the nonlinear solver error and explicit treatment of nonlocal effective field contributions introduce slight perturbations to this discrete energy identity.

The tangent plane scheme introduces artificial dissipation and only provides a discrete

energy estimate. Precisely, on meshes satisfying the angle condition (1.25), the projection update $\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \boldsymbol{\mathcal{I}}_h(\hat{\boldsymbol{m}}_h^{i+1}/|\hat{\boldsymbol{m}}_h^{i+1}|)$ does not increase the energy (and is energy decreasing in general).

Analogously to the tangent plane scheme, the first-order integrator from [KW18] satisfies a discrete energy estimate on meshes satisfying the angle condition (1.25). For the second-order scheme proposed in [KW18], although our numerical experiments in Chapter 5 support its practical applicability, the stability of the scheme remains open.

In the remainder of this section, we present the original versions of the three (families) of schemes. For the sake of clarity, here we restrict the presentation to the exchange-only case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \boldsymbol{\Delta m}$. In Section 1.6 below, as well as in the works presented in Chapter 2–6, the (extended) integrators are presented in such a way that more general effective fields are covered. For any of the presented schemes, we consider a triangulation $\mathcal{T}_h$ as defined in Section 1.4, and for some fixed $J \in \mathbb{N}$, we define the time-step size $k := T/J > 0$.

1.5.2 Midpoint scheme

The following algorithm states the implicit midpoint scheme (MPS) proposed by Bartels & Prohl in [BP06].

Algorithm 1.5.1 (MPS, [BP06]). **Input:** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$.
Loop: *For all time-steps $i = 0, \ldots, J-1$, compute $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that*

$$\langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h \rangle_h = -\ell_{\text{ex}}^2 \langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{\Delta}_h \boldsymbol{m}_h^{i+1/2}, \boldsymbol{w}_h \rangle_h + \alpha \langle \boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h \rangle_h. \quad (1.27)$$

Output: *Sequence of approximations $\{\boldsymbol{m}_h^i\}_{i=0}^J$.*

The discrete variational formulation (1.27) is based on the Gilbert form $(\mathcal{G})$ of LLG. The fundamental ingredients of the scheme are the midpoint rule for the time discretization, the mass-lumped scalar product $\langle \cdot, \cdot \rangle_h$, and the discrete Laplace operator (1.24). Utilization of the mass-lumped integration rule $\langle \cdot, \cdot \rangle_h$ preserves nodewise unit-length $|\boldsymbol{m}_h^{i+1}(\boldsymbol{z})| = |\boldsymbol{m}_h^i(\boldsymbol{z})|$ for any $\boldsymbol{z} \in \mathcal{N}_h$, which is immediately obtained by testing with $\boldsymbol{w}_h := \boldsymbol{m}_h^{i+1/2}(\boldsymbol{z})\phi_{\boldsymbol{z}}$. Moreover, testing (1.27) with $\boldsymbol{w}_h := \alpha d_t \boldsymbol{m}_h^{i+1} - \ell_{\text{ex}}^2 \boldsymbol{\Delta}_h \boldsymbol{m}_h^{i+1/2}$ and summation over $i = 0, \ldots, J-1$ reveals that Algorithm 1.5.1 realizes a discrete energy equality

$$\frac{\ell_{\text{ex}}^2}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^J\|_{\boldsymbol{L}^2(\Omega)}^2 + \alpha k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_h^{i+1}\|_h^2 = \frac{\ell_{\text{ex}}^2}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2, \quad (1.28)$$

resembling the continuous energy law (1.18). The MPS is unconditionally convergent towards a weak solution of LLG and formally of second order in time, but it requires the solution of the nonlinear system (1.27) of equations in each time-step. Existence of solutions of the discrete variational formulation (1.27) can be shown using the Brouwer fixed-point theorem; see, e.g., [PRS18b, Proposition 3]. Existence and uniqueness of a solution can be obtained by means of the Banach fixed-point theorem, but require the strong CFL condition $k = o(h^2)$; see [BP06, Lemma 4.1]. Linearization strategies for the nonlinear system (1.27), necessary

to arrive at a practically applicable scheme, are discussed in Section 1.6.2. We note that the nonlinearity is a consequence of the first term on the right-hand side of (1.27). The second term $\boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}$ on the right-hand side, at first glance also nonlinear in $\boldsymbol{m}_h^{i+1}$, can equivalently be rewritten, e.g., as $k^{-1} \boldsymbol{m}_h^i \times \boldsymbol{m}_h^{i+1}$.

1.5.3 Tangent plane scheme

In the mathematical literature, a fully explicit version of the tangent plane scheme was already proposed by Alouges & Jaisson in [AJ06] and analyzed in [BKP08]. Here, we present the semi-implicit tangent plane scheme from [Alo08] that covers the original version from [AJ06] with the choice $\theta = 0$. In the following algorithm, we state the first-order tangent plane scheme (TPS1) proposed and analyzed by Alouges in [Alo08].

Algorithm 1.5.2 (TPS1, [Alo08])**.** **Input:** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ and $0 \le \theta \le 1$.
Loop: *For all time-steps* $i = 0, \dots, J-1$*, iterate:*

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

$$\alpha \langle \boldsymbol{v}_h^i, \boldsymbol{\varphi}_h \rangle + \langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^i, \boldsymbol{\varphi}_h \rangle + \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{\nabla} \boldsymbol{v}_h^i, \boldsymbol{\nabla} \boldsymbol{\varphi}_h \rangle = -\ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^i, \boldsymbol{\nabla} \boldsymbol{\varphi}_h \rangle . \qquad (1.29)$$

(ii) *Define* $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ *by*

$$\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^i(\boldsymbol{z}) + k \boldsymbol{v}_h^i(\boldsymbol{z})}{|\boldsymbol{m}_h^i(\boldsymbol{z}) + k \boldsymbol{v}_h^i(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h. \qquad (1.30)$$

Output: *Sequence of discrete functions* $\left\{ (\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1}) \right\}_{i=0}^{J-1}$.

As for the classical θ-method for the heat equation, the parameter $\theta \in [0,1]$ in (1.29) modulates the 'degree of implicitness' of the scheme: The choice $\theta = 0$ leads to a fully explicit scheme (explicit Euler scheme); the choice $\theta = 1/2$ is a Crank–Nicolson-type method, which in the case of the heat equation is of second-order in time; the choice $\theta = 1$ is a sort of implicit Euler method. The discrete variational formulation (1.29) of step (i) is a weak formulation of the alternative form $(\mathcal{A})$ of LLG posed in the discrete tangent space $\boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^i]$. It provides an approximation $\boldsymbol{v}_h^i \approx \partial_t \boldsymbol{m}(t_i)$ of the time derivative, which is then used for the time-stepping of step (ii). To ensure that the magnetization belongs to the set of admissible discrete magnetizations $\boldsymbol{\mathcal{M}}_h$, the time-stepping in (1.30) employs the nodal projection.

The scheme requires the solution of only one (constrained, but always well-posed) linear system (1.29) per time-step, and is formally first-order accurate in time for any $0 \le \theta \le 1$. In contrast to the classical θ-method for the heat equation, however, the choice $\theta = 1/2$ does not lead to any improvement of the first-order convergence in time, because of the tangent plane constraint $\boldsymbol{v}_h^i \in \boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^i]$ and the presence of the nodal projection in (1.30).

Convergence towards a weak solution for the original version ($\theta = 0$) was proved in [AJ06] first passing the time-step size to the limit $k \to 0$, and only then separately taking the limit $h \to 0$ of the mesh-size, i.e., their analysis imposes a not precisely quantified coupling of the discretization parameters to obtain convergence. Later in [BKP08], the necessary coupling

condition was quantified as $k = o(h^{5/2})$. Finally in [Alo08], the algorithm of [AJ06] is generalized to Algorithm 1.5.2. There, imposing the angle-condition (1.25) from [Bar05] on the family of meshes $\{\mathcal{T}_h\}_{h>0}$, a further improvement of the necessary coupling is established, namely, the coupling $k = o(h^2)$ for all $0 \leq \theta < 1/2$, as well as the coupling condition $k = o(h)$ for $\theta = 1/2$, and unconditional convergence as $h, k \to 0$ for $1/2 < \theta \leq 1$.

1.5.4 Recent predictor-corrector scheme

Only recently in [KW18], Kim & Wilkening proposed a finite element-based numerical scheme discretizing the Landau–Lifshitz form $(\mathcal{LL})$. We note that in the mathematical literature this is not the first such integrator: Already in [Cim09], similar to the midpoint scheme, the author proposes a mass-lumped integrator based on the Landau–Lifshitz form $(\mathcal{LL})$ and the implicit midpoint rule in time. We state the first-order predictor-corrector method (`PC1`) from [KW18].

Algorithm 1.5.3 (`PC1`, [KW18])**.** **Input:** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ and $0 \leq \theta \leq 1$.
Loop: *For all time-steps* $i = 0, \ldots, J-1$*, iterate:*

(i) *Compute* $\boldsymbol{v}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$*, it holds that*

$$\begin{aligned}
(1 + \alpha^2)\langle \boldsymbol{v}_h^i, \boldsymbol{w}_h \rangle_h = &-\ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^i \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i), \boldsymbol{w}_h \rangle_h \\
&- \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^i \times (\boldsymbol{m}_h^i \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i)), \boldsymbol{w}_h \rangle_h\,.
\end{aligned} \tag{1.31}$$

(ii) *Define* $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ *by*

$$\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})}{|\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \textit{for all } \boldsymbol{z} \in \mathcal{N}_h. \tag{1.32}$$

Output: *Sequence of discrete functions* $\big\{(\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1})\big\}_{i=0}^{J-1}$.

In the same work [KW18], also a second-order variant of Algorithm 1.5.3 is proposed, which we will present in Section 1.6 below together with implicit-explicit (IMEX) extensions of these schemes.

As for the tangent plane scheme, the parameter $0 \leq \theta \leq 1$ modulates the implicitness of the scheme, the unknown $\boldsymbol{v}_h^i$ approximates the time derivative $\partial_t \boldsymbol{m}(t_i)$, and the update is defined via nodal projection. Differently to the tangent plane scheme, Algorithm 1.5.3 relies on the mass-lumped integration in space, and is originally formulated to integrate (nonlocal) lower-order effective field contributions implicitly in time. In view of Remark 5.3.4 in Chapter 5, it is not surprising that for Algorithm 1.5.3 a convergence result analogous to the one for the first-order tangent plane scheme from [Alo08] is obtained, i.e., unconditional convergence for $1/2 < \theta \leq 1$, weak coupling $k = o(h)$ to guarantee convergence for $\theta = 1/2$, and strong coupling $k = o(h^2)$ to guarantee convergence for $0 \leq \theta < 1/2$. We note that the formulation of the main theorem [KW18, Theorem 2.2] (erroneously) does not require any CFL condition for $\theta = 1/2$. This is due to an inaccuracy in their proof when passing the inequality [KW18, (4.62)] to the limit $k, h \to 0$. In our analysis in Section 5.3.4, we cure

the proof of the theorem by tracing the stability of `PC1` back to arguments from [Alo08] for the tangent plane scheme.

This is an appropriate point in this work to mention one of the new theoretical contributions of this thesis: While the original work [KW18] lacks any discussion of the non-obvious well-posedness of Algorithm 1.5.3, in the preprint [MPPR21] of our work presented in Chapter 5 we show unconditional well-posedness of Algorithm 1.5.3. Precisely, the system (1.31) always admits a unique solution $\boldsymbol{v}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$, for all discretization parameters $h, k > 0$, for any parameter of implicitness $0 \leq \theta \leq 1$, and — surprisingly — for any Gilbert damping parameter $\alpha \geq 0$ including the limit case $\alpha = 0$.

1.6 Extensions of the original integrators

With a fully discrete, well-posed, and stable scheme at hand, usually further steps are necessary to arrive not only at a practically applicable integrator, but at a computationally feasible one.

1.6.1 Practical requirements

Taking into account that computational resources — hardware and time — are limited, certain aspects of a numerical scheme have to be considered with special care.

- Nonlocal effective field contributions should not, if by any means possible, be integrated implicitly in time. This is because their implicit treatment leads to infeasible computational costs, i.e., either inducing a densely populated (linear) system of equations, or relying on a costly inner iteration in the iterative (linear or nonlinear) solver.

- Higher-order (in time) accurate schemes are to be preferred over first-order schemes, if their computational costs are roughly comparable. We note that one has to be careful not to spoil the formal second-order accuracy by explicit in-time integration of (nonlocal) effective field contributions.

- Integrators requiring weaker coupling of the discretization parameters — usually the time-step size k is restricted by the mesh size h — are to be preferred over those requiring stronger coupling. This allows for coarser time discretizations, and hence, fewer time-steps resulting in faster simulations.

- As the space discretization leads to linear systems of dimension scaling proportionally to $1/h^3$, direct solvers are only of limited use for the schemes presented in Section 1.5 in practice. Hence, for resource-efficient simulations, the arising (non-symmetric) linear systems are numerically solved by iterative solvers like, e.g., GMRES [SS86]. Appropriate preconditioners can drastically reduce the number of necessary iterations to reach a certain accuracy, and hence, reduce computational time.

The possibility to choose from multiple numerical schemes is always an advantage, as the following brief discussion of two state-of-the-art numerical integrators for LLG motivates. Even with respect to the weakened discrete requirements, none of the integrators available in the literature is perfect: For example, the tangent plane scheme variant from [AKST14]

realizes the discrete unit-length constraint, is formally second-order accurate in time, and imposes only a weak coupling of the discretization parameters. With a further modification from [DPP+20], it can even treat nonlocal effective field contributions explicitly in time. Stability of this integrator, however, is theoretically only guaranteed for certain types of triangulations $\mathcal{T}_h$. Moreover, artificial damping is introduced to the system, as the theory only provides a discrete energy inequality.

Another example is the (ideal) midpoint scheme proposed in [BP06], which realizes a discrete energy equality, preserves discrete unit-length, is formally second-order accurate in time, and allows for more general triangulations. Nonlocal effective field contributions can be integrated explicitly in time [PRS18b]. On the downside, a practical realization of the midpoint scheme has to linearize the arising nonlinear discrete system, and hence, requires an inner iteration in each time-step. Moreover, any linearization suggested in the literature requires a strong coupling of the discretization parameters $k = o(h^2)$ to ensure convergence of the nonlinear solver.

In the remainder of this section, we present modifications and extensions of the three original schemes from Section 1.5. Any variant aims to improve the practical feasibility of the corresponding original scheme. As often in mathematics and life in general, usually nothing comes for free and an improvement of one aspect is made possible by accepting a (slight) drawback in another.

To clearly and adequately present the intention of the various modified schemes below, we consider the more general effective field $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$. Here, $\boldsymbol{f} \in C^1([0,T]; \boldsymbol{L}^2(\Omega))$ denotes some applied external field, and $\boldsymbol{\pi} \colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ denotes a linear, self-adjoint, and bounded operator collecting lower-order effective field contributions like, e.g., uniaxial anisotropy (1.9a) or the nonlocal stray field (1.11a). Moreover, let $\boldsymbol{\pi}_h \colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathcal{S}^1(\mathcal{T}_h)^3$ and $\boldsymbol{f}_h^i$ denote some (here not further specified) approximations of the continuous operator $\boldsymbol{\pi}$ and data $\boldsymbol{f}(t_i)$, respectively. For precise requirements on the approximations $\boldsymbol{\pi}_h$ and $\boldsymbol{f}_h$, we refer to the theorems in Section 2–6.

We note that this setting covers all standard energy contributions, i.e., all energy contributions introduced in Section 1.2, with exception of the non-standard DMI energy contribution (1.14a). A rigorous extension of the family of midpoint schemes to cover DMI is proposed in Chapter 3 of this thesis. The family of tangent plane integrators has been extended to DMI in the joint work [HPP+19], which is presented in Chapter 2. An extension of the family of predictor-corrector methods from [KW18] to cover DMI is still missing in the literature, but is expected to be possible, in particular not requiring essentially different tools from those used in Chapter 2 and Chapter 3.

1.6.2 Extensions of the midpoint scheme

Implicit-explicit midpoint scheme

In [PRS18b], the authors apply a modification to Algorithm 1.5.1 in order to overcome one of its practical drawbacks, the implicit integration of (nonlocal) lower-order effective field contributions $\boldsymbol{\pi}_h(\boldsymbol{m}_h^{i+1/2})$, while preserving the second-order accuracy of MPS. The underlying idea is to apply a two-step Adams–Bashforth method in combination with linearity of

$\boldsymbol{\pi}_h$ to approximate

$$\boldsymbol{\pi}_h(\boldsymbol{m}_h^{i+1/2}) \approx \boldsymbol{\Pi}_h(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) := (3/2)\boldsymbol{\pi}_h(\boldsymbol{m}_h^i) - (1/2)\boldsymbol{\pi}_h(\boldsymbol{m}_h^{i-1}).$$

Application of this approximation to Algorithm 1.5.1 results in the following second-order implicit-explicit midpoint scheme (`MPS+IMEX`).

Algorithm 1.6.1 (`MPS+IMEX`, [PRS18b]). ***Input:*** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$.
Loop: *For all time-steps* $i = 0, \ldots, J-1$*, compute* $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$*, it holds that*

$$\begin{aligned}
\langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h \rangle_h = {}& -\langle \boldsymbol{m}_h^{i+1/2} \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h \boldsymbol{m}_h^{i+1/2} + \boldsymbol{\Pi}_h(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) + \boldsymbol{f}_h^{i+1/2}], \boldsymbol{w}_h \rangle_h \\
& + \alpha \langle \boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h \rangle_h
\end{aligned} \tag{1.33}$$

Output: *Sequence of approximations* $\{\boldsymbol{m}_h^i\}_{i=0}^J$.

The implicit-explicit scheme improves the original integrator `MPS` in a practical sense as it requires only one evaluation of $\boldsymbol{\pi}_h(\boldsymbol{m}_h^i)$ per time-step. We note that the discrete system is still nonlinear, as the highest-order exchange contribution is still treated implicitly in time in order to preserve unconditional stability of the scheme. A nonlinear solver for `MPS+IMEX`, however, is considerably cheaper than for `MPS`, since inside the solver no expensive stray field evaluations are necessary for `MPS+IMEX`. Due to the second-order accuracy of the Adams–Bashforth method, the second-order accuracy of the original midpoint scheme (Algorithm 1.5.1), is (heuristically and empirically) preserved. A drawback of this implicit-explicit approximation is a slight perturbation of the discrete energy equality (1.28) realized by the ideal midpoint scheme (Algorithm 1.5.1). Anyhow, we note that also apart from the IMEX approach a perturbation of (1.28) is unavoidable in practice, as linearization of either, (1.27) or (1.33), introduces a perturbation of (1.28).

Practical implicit-explicit midpoint schemes

Both, Algorithm 1.5.1 and Algorithm 1.6.1, require the solution of a nonlinear system of equations in each time-step. In the original work [BP06], the authors propose a fixed-point iteration for the approximate solution of the discrete nonlinear equation, which is well-defined if the discretization parameters satisfy the strong coupling condition $k = o(h^2)$. Then, the fixed-point iteration converges towards the solution of the nonlinear system, which is unique by the Banach fixed-point theorem. The fixed-point iteration from [BP06] does not preserve the discrete unit-length constraint, but the authors prove that the discrete magnetization length remains uniformly bounded.

In [Bar06] a constraint preserving fixed-point iteration for the midpoint scheme is proposed, which is shown to be convergent under the same coupling condition $k = o(h^2)$.

Based on the Newton scheme, in [BBNP14, Section 1.4.1] the authors employ a linearization of the nonlinear system of the midpoint scheme. Their 2D numerical experiments give hope for a less restrictive CFL condition than for the fixed-point iterations from [BP06] and [Bar06].

In an ongoing work presented in Chapter 3 of this thesis, for micromagnetic simulations in 3D and considering a general effective field, we apply Newton's method and a

constraint preserving fixed-point iteration, respectively, to the implicit-explicit nonlinear midpoint scheme `MPS+IMEX` to obtain two practical versions of the midpoint scheme. Our well-posedness and convergence analysis leads to the same coupling condition $k = o(h^2)$ for the constraint preserving fixed-point iteration, while it requires the even stronger condition $k = o(h^{7/3})$ to ensure convergence of the Newton linearization. Experimentally, both linearization strategies require the same coupling condition $k = o(h^2)$, qualitatively and also quantitatively, i.e., with the same hidden constant. However, the experiments show that in terms of iteration numbers required to achieve a prescribed tolerance, the Newton linearization clearly outperforms the fixed-point iteration.

Finally, in a fully practical adaption of the original `MPS` from Section 1.5.2, two sources of perturbation are introduced to the discrete energy identity (1.28); see (3.17), (3.28), and (3.33) in Chapter 3 for a precise quantification. Moreover, depending on the linearization strategy, discrete unit-length might only be conserved asymptotically; see Theorem 3.3.9(i)–(ii) in Chapter 3 for the practical midpoint scheme with Newton linearization.

1.6.3 Extensions of the first-order tangent plane scheme

In [AKT12], and independently in [Gol12], Algorithm 1.5.2 is extended to cover the more general effective field $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$ as considered in this section. Both works integrate lower-order effective field contributions explicitly in time. In particular, [Gol12] covers the approximate computation of lower-order terms, including the stray field, which is inevitable in practice. Here, we do not explicitly state this generalization of Algorithm 1.5.2, but note that it can be obtained from combining Algorithm 1.6.2(i) and Algorithm 1.5.2(ii).

Projection-free first-order tangent plane scheme

In the following algorithm, we state a projection-free variant of the first-order tangent plane scheme (`TPS1PF`), which was proposed in [AHP+14] adapting original ideas from [Bar16] to LLG. Analogously to the original first-order tangent plane scheme from [Alo08], first an approximate $\boldsymbol{v}_h^i \in \mathcal{K}_h[\boldsymbol{m}_h^i]$ of the time derivative is computed. In contrast to `TPS1`, when advancing $\boldsymbol{m}_h^i \rightsquigarrow \boldsymbol{m}_h^{i+1}$ the nodal projection in the update step is omitted. At the price of losing nodal unit-length, i.e., here $\boldsymbol{m}_h^i \notin \mathcal{M}_h$ in general, this modification leads to a linear update formula.

Algorithm 1.6.2 (`TPS1PF`, [AHP+14]). ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ *and* $0 \leq \theta \leq 1$.
Loop: *For all time-steps* $i = 0, \ldots, J-1$, *iterate:*

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

$$\alpha \langle \boldsymbol{v}_h^i, \boldsymbol{\varphi}_h \rangle + \langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^i, \boldsymbol{\varphi}_h \rangle + \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{\nabla v}_h^i, \boldsymbol{\nabla \varphi}_h \rangle$$
$$= -\ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla m}_h^i, \boldsymbol{\nabla \varphi}_h \rangle + \langle \boldsymbol{\pi}_h(\boldsymbol{m}_h^i), \boldsymbol{\varphi}_h \rangle + \langle \boldsymbol{f}_h^i, \boldsymbol{\varphi}_h \rangle \,.$$

(ii) *Define* $\boldsymbol{m}_h^{i+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$ *by*
$$\boldsymbol{m}_h^{i+1} := \boldsymbol{m}_h^i + k \boldsymbol{v}_h^i.$$

Output: *Sequence of discrete functions* $\left\{ (\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1}) \right\}_{i=0}^{J-1}$.

First of all, we note that for $\boldsymbol{m}_h^i \in \mathcal{S}^1(\mathcal{T}_h)$ the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^i]$ can be defined analogously to (1.21) as long as $\boldsymbol{m}_h^i(\boldsymbol{z}) \neq \boldsymbol{0}$ for all nodes $\boldsymbol{z} \in \mathcal{N}_h$. The analysis of TPS1PF even shows $|\boldsymbol{m}_h^i(\boldsymbol{z})| \geq 1$ for all nodes $\boldsymbol{z} \in \mathcal{N}_h$ and all time-steps $i = 0, \ldots, J-1$. Hence, in step (i) the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^i]$ is always well-defined.

Although nodewise unit-length is lost, the analysis of TPS1PF reveals that the constraint violation is uniformly controlled by the time-step size $k > 0$. The convergence result and the formal order in time of TPS1 are preserved, and additionally, the analysis of TPS1PF does not impose an angle condition (1.25) on the family of meshes $\{\mathcal{T}_h\}_{h>0}$. Moreover, only recently in [FT17a], the authors were able to prove strong convergence of TPS1PF, and even establish the expected convergence rates as a priori error estimates.

Almost second-order tangent plane scheme

We state the almost second-order tangent plane scheme (TPS2) proposed and analyzed by Alouges et al. in [AKST14]. The underlying idea was already mentioned in [AKT12] and is motivated by a Taylor expansion of the projection-update (1.30) with respect to k, yielding that the extrapolation

$$\boldsymbol{m}(t+k) \approx \frac{\boldsymbol{m}(t) + k\boldsymbol{v}}{|\boldsymbol{m}(t) + k\boldsymbol{v}|} \quad \text{with} \quad \boldsymbol{v} := \partial_t \boldsymbol{m}(t) + \frac{k}{2}\big(\partial_{tt}\boldsymbol{m}(t) - (\boldsymbol{m}(t) \cdot \partial_{tt}\boldsymbol{m}(t))\boldsymbol{m}(t)\big)$$

is third-order accurate $\mathcal{O}(k^3)$ in the time-step size $k > 0$. Then, again from the alternative form $(\mathcal{A})$ of LLG a discrete system is derived, defining

$$\partial_t \boldsymbol{m}(t_i) + \frac{k}{2}\big(\partial_{tt}\boldsymbol{m}(t_i) - (\boldsymbol{m}(t_i) \cdot \partial_{tt}\boldsymbol{m}(t_i))\boldsymbol{m}(t_i)\big) \approx \boldsymbol{v}_h^i \in \mathcal{K}_h[\boldsymbol{m}_h^i] \,.$$

Defining the notations

$$\rho(k) := |k \log k| \qquad \text{and} \qquad \widetilde{W}_k(s) := \begin{cases} \alpha + \frac{k}{2}\min\{s, B\} & \text{for } s \geq 0, \\ \frac{2\alpha^2}{2\alpha + k\min\{-s, B\}} & \text{for } s < 0, \end{cases}$$

with some cut-off $B \in \mathbb{R}^+$ and $W_k(\boldsymbol{m}_h) := \widetilde{W}_k(\boldsymbol{m}_h \cdot \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_h))$, they arrive at the following second-order tangent plane scheme (TPS2).

Algorithm 1.6.3 (TPS2, [AKST14]). ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: *For all time-steps $i = 0, \ldots, J-1$, iterate:*

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

$$\langle W_k(\boldsymbol{m}_h^i)\boldsymbol{v}_h^i, \boldsymbol{\varphi}_h\rangle + \langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^i, \boldsymbol{\varphi}_h\rangle + \frac{\ell_{\text{ex}}^2}{2}k[1 + \rho(k)]\langle \boldsymbol{\nabla}\boldsymbol{v}_h^i, \boldsymbol{\nabla}\boldsymbol{\varphi}_h\rangle - \frac{k}{2}\langle \boldsymbol{\pi}_h(\boldsymbol{v}_h^i), \boldsymbol{\varphi}_h\rangle$$
$$= -\ell_{\text{ex}}^2\langle \boldsymbol{\nabla}\boldsymbol{m}_h^i, \boldsymbol{\nabla}\boldsymbol{\varphi}_h\rangle + \langle \boldsymbol{\pi}_h(\boldsymbol{m}_h^i), \boldsymbol{\varphi}_h\rangle + \langle \boldsymbol{f}_h^{i+1/2}, \boldsymbol{\varphi}_h\rangle \,. \tag{1.34}$$

(ii) *Define $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$ by*

$$\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})}{|\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h. \tag{1.35}$$

Output: *Sequence of discrete functions* $\left\{(\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1})\right\}_{i=0}^{J-1}$.

Like in its first-order counterpart, Algorithm 1.5.2, the variational formulation (1.34) is posed in the discrete tangent space $\boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^i]$, and the update (1.35) employs the nodal projection to enforce the pointwise constraint at the nodes of the triangulation. The Gilbert damping parameter α is replaced by the weight function $W_k(\boldsymbol{m}_h^i)$, which gives rise to the second-order of the scheme and ensures coercivity of the system by use of the cut-off parameter $B \in \mathbb{R}^+$. The Crank–Nicolson type choice $\theta = 1/2$ is fixed in TPS2, with an additional stabilization term $\rho(k)$ added to the highest-order exchange contribution. This additional stabilization allows to prove unconditional convergence of TPS2 at the cost of reducing the formal order of the scheme from full second-order to 'almost second-order', i.e., for any $\varepsilon > 0$ the rate $2 - \varepsilon$ is realized as $k \to 0$. Assuming the angle condition (1.25) on the family of triangulations $\{\mathcal{T}_h\}_{h>0}$, the method is unconditionally convergent towards a weak solution of LLG and formally of almost second-order in time. We note that there is some flexibility in the definition of $W_k(\cdot)$ and $\rho(k)$, which we describe in more detail in [DPP$^+$20, (2.7)], where we additionally propose a natural choice to avoid the additional parameter $B \in \mathbb{R}^+$. In particular, the choice $\rho(k) := 0$ formally leads to full second-order in-time accuracy at the cost of imposing the weak coupling condition $k = o(h)$ to obtain stability of the scheme; see [AKST14, Theorem 2] and [HPP$^+$19, Remark 3(ii)–(iii)].

A considerable drawback of TPS2, however, is its implicit integration of lower-order effective field contributions $\boldsymbol{\pi}_h(\boldsymbol{m}_h^i + (k/2)\boldsymbol{v}_h^i)$, in particular including the nonlocal stray field. Despite the linearity of (1.34), practically this has to be dealt with by a costly inner iteration in the (linear) solver.

Almost second-order implicit-explicit tangent plane scheme

In [DPP$^+$20], we applied another modification to Algorithm 1.6.3 in order to overcome its practical drawback, namely the implicit integration of (nonlocal) lower-order effective field contributions $\boldsymbol{\pi}_h(\boldsymbol{m}_h^i + (k/2)\boldsymbol{v}_h^i)$, while preserving the almost second-order accuracy of TPS2. The underlying idea is to apply a two-step Adams–Bashforth method in combination with linearity of $\boldsymbol{\pi}_h$ to approximate

$$\boldsymbol{\pi}_h(\boldsymbol{m}_h^i + (k/2)\boldsymbol{v}_h^i) \approx \boldsymbol{\Pi}_h(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) := (3/2)\boldsymbol{\pi}_h(\boldsymbol{m}_h^i) - (1/2)\boldsymbol{\pi}_h(\boldsymbol{m}_h^{i-1}).$$

With the definitions of $W_k(\cdot)$ and $\rho(k)$ from Algorithm 1.6.3, applying the approximation to TPS2 results in the following almost second-order implicit-explicit tangent plane scheme (TPS2+IMEX).

Algorithm 1.6.4 (TPS2+IMEX, [DPP$^+$20]). **Input:** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$. *Define* $\boldsymbol{m}_h^{-1} := \boldsymbol{m}_h^0$. **Loop:** *For all time-steps* $i = 0, \ldots, J-1$*, iterate:*

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

$$\langle W_k(\boldsymbol{m}_h^i)\boldsymbol{v}_h^i, \boldsymbol{\varphi}_h \rangle + \langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^i, \boldsymbol{\varphi}_h \rangle + \frac{\ell_{\mathrm{ex}}^2}{2}k[1 + \rho(k)]\langle \boldsymbol{\nabla}\boldsymbol{v}_h^i, \boldsymbol{\nabla}\boldsymbol{\varphi}_h \rangle$$

$$= -\ell_{\mathrm{ex}}^2\langle \boldsymbol{\nabla}\boldsymbol{m}_h^i, \boldsymbol{\nabla}\boldsymbol{\varphi}_h \rangle + \langle \boldsymbol{\Pi}_h(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}), \boldsymbol{\varphi}_h \rangle + \langle \boldsymbol{f}_h^{i+1/2}, \boldsymbol{\varphi}_h \rangle. \tag{1.36}$$

(ii) Define $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ by

$$\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})}{|\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \textit{for all } \boldsymbol{z} \in \mathcal{N}_h. \tag{1.37}$$

Output: *Sequence of discrete functions $\left\{(\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1})\right\}_{i=0}^{J-1}$.*

The scheme inherits all analytical properties from `TPS2`, and requires only one evaluation of $\boldsymbol{\pi}_h(\boldsymbol{m}_h^i)$ per time-step. Hence, besides the evaluation of $\boldsymbol{\pi}_h(\boldsymbol{m}_h^i)$, the scheme requires the solution of only one (constrained, but always well-posed) linear system per time-step. The method is almost second-order accurate in time, and is computationally only slightly more expensive than the first-order tangent plane scheme `TPS1`, because the mass term needs to be updated at each time-step.

1.6.4 Extensions of the recent predictor-corrector scheme

Second-order predictor-corrector scheme `PC2`

In [KW18], the authors propose a second-order variant `PC2` of their first-order predictor-corrector scheme `PC1`. In `PC2`, the predictor step of Algorithm 1.5.3 remains unchanged, computing an approximation $\boldsymbol{v}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$ of the time derivative $\partial_t \boldsymbol{m}(t_i)$. The projection update of `PC1` is replaced by a linear system of equations defining $\boldsymbol{m}_h^{i+1}$. This linear system for the corrector is also a discretization of the Landau–Lifshitz form $(\mathcal{LL})$, where the effective field is explicitly evaluated on the predictor $\boldsymbol{m}_h^i + (k/2)\boldsymbol{v}_h^i$. Using the cross product structure of $(\mathcal{LL})$ in combination with the mass-lumped integration rule, the remaining terms are approximated such that the resulting discrete system is linear and discrete unit-length $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ is ensured. The obtained algorithm reads as follows.

Algorithm 1.6.5 (`PC2`, [KW18]). **Input:** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ *and* $0 \le \theta \le 1$.
Loop: *For all time-steps $i = 1, \ldots, J-1$, iterate:*

(i) Compute $\boldsymbol{v}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that

$$\begin{aligned}
(1+\alpha^2)\langle \boldsymbol{v}_h^i, \boldsymbol{w}_h \rangle_h &= -\langle \boldsymbol{m}_h^i \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{\pi}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{f}_h^{i+\theta}], \boldsymbol{w}_h \rangle_h \\
&\quad - \alpha \langle \boldsymbol{m}_h^i \times \left(\boldsymbol{m}_h^i \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{\pi}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{f}_h^{i+\theta}]\right), \boldsymbol{w}_h \rangle_h.
\end{aligned}$$

(ii) Define $\hat{\boldsymbol{m}}_h^{i+1/2} := \boldsymbol{m}_h^i + (k/2)\boldsymbol{v}_h^i$ and compute $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that

$$\begin{aligned}
(1+\alpha^2)\langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h \rangle_h &= -\langle \boldsymbol{m}_h^{i+1/2} \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h \hat{\boldsymbol{m}}_h^{i+1/2} + \boldsymbol{\pi}_h(\hat{\boldsymbol{m}}_h^{i+1/2}) + \boldsymbol{f}_h^{i+1/2}], \boldsymbol{w}_h \rangle_h \\
&\quad - \alpha \langle \boldsymbol{m}_h^{i+1/2} \times \left(\hat{\boldsymbol{m}}_h^{i+1/2} \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\hat{\boldsymbol{m}}_h^{i+1/2}) + \boldsymbol{\pi}_h(\hat{\boldsymbol{m}}_h^{i+1/2}) + \boldsymbol{f}_h^{i+1/2}]\right), \boldsymbol{w}_h \rangle_h.
\end{aligned}$$

Output: *Sequence of discrete functions $\left\{(\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1})\right\}_{i=0}^{J-1}$.*

The numerical experiments in [KW18] show second-order accuracy of the scheme. However, no well-posedness and no convergence analysis is given in [KW18]. In the recent work [MPPR21] presented in Chapter 5 of this thesis, we prove unconditional well-posedness of Algorithm 1.6.5, again including the limit case $\alpha = 0$. Further, our well-posedness analysis extends to the case where $\boldsymbol{m}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$, but not necessarily $\boldsymbol{m}_h^i \notin \boldsymbol{\mathcal{M}}_h$. This is necessary, since $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ only holds in theory as it is implicitly derived from the corrector system in Algorithm 1.6.5(ii), which, in practice, is approximately solved by use of an inexact (iterative) solver. The stability analysis of PC2, however, still remains open.

Implicit-explicit first- and second-order predictor-corrector schemes

Both integrators from [KW18], PC1 and PC2, integrate (nonlocal) lower-order effective field contributions $\boldsymbol{\pi}_h(\cdot)$ implicitly in time, restricting their practical feasibility. In the recent work [MPPR21] presented in Chapter 5, we also apply a modification to Algorithm 1.5.3 and Algorithm 1.6.5, respectively, in order to overcome their practical drawback, while preserving their formal first- and second-order accuracy. Using the backward difference quotient and using linearity of $\boldsymbol{\pi}_h$ yields the approximation

$$\boldsymbol{\pi}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) \approx \boldsymbol{\Pi}_h^\theta(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) := (1+\theta)\boldsymbol{\pi}_h(\boldsymbol{m}_h^i) - \theta \boldsymbol{\pi}_h(\boldsymbol{m}_h^{i-1}).$$

Application to Algorithm 1.5.3 and Algorithm 1.6.5, respectively, results in the implicit-explicit predictor-corrector methods PC1+IMEX and PC2+IMEX.

Algorithm 1.6.6 (PC1+IMEX, [MPPR21]). ***Input:*** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ *and* $0 \leq \theta \leq 1$.
Define $\boldsymbol{m}_h^{-1} := \boldsymbol{m}_h^0$.
Loop: *For all time-steps* $i = 0, \ldots, J-1$, *iterate:*

(i) *Compute* $\boldsymbol{v}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$(1+\alpha^2)\langle \boldsymbol{v}_h^i, \boldsymbol{w}_h \rangle_h = -\langle \boldsymbol{m}_h^i \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{\Pi}_h^\theta(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) + \boldsymbol{f}_h^{i+\theta}], \boldsymbol{w}_h \rangle_h$$
$$- \alpha \langle \boldsymbol{m}_h^i \times \left(\boldsymbol{m}_h^i \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{\Pi}_h^\theta(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) + \boldsymbol{f}_h^{i+\theta}] \right), \boldsymbol{w}_h \rangle_h.$$

(ii) *Define* $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ *by*

$$\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^i(\boldsymbol{z}) + k \boldsymbol{v}_h^i(\boldsymbol{z})}{|\boldsymbol{m}_h^i(\boldsymbol{z}) + k \boldsymbol{v}_h^i(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h.$$

Output: *Sequence of discrete functions* $\left\{ (\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1}) \right\}_{i=0}^{J-1}$.

We note that (a slightly different version of) PC1+IMEX can be obtained from PC1 by the approximation $k \boldsymbol{\pi}_h(\boldsymbol{v}_h^i) \approx \boldsymbol{0}$. This would not deteriorate the order of the scheme, as PC1 is already restricted to first-order in time. Our analysis in Chapter 5 shows that PC1+IMEX inherits all analytical properties of PC1. Practically, PC1+IMEX is advantageous, as besides the evaluation of $\boldsymbol{\pi}_h(\boldsymbol{m}_h^i)$, it requires the solution of only one sparse linear system per time-step.

Algorithm 1.6.7 (PC2+IMEX, [MPPR21])**.** ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ *and* $0 \le \theta \le 1$.
Define $\boldsymbol{m}_h^{-1} := \boldsymbol{m}_h^0$.
Loop: *For all time-steps* $i = 0, \dots, J-1$, *iterate:*

(i) *Compute* $\boldsymbol{v}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$(1+\alpha^2)\langle \boldsymbol{v}_h^i, \boldsymbol{w}_h \rangle_h = -\langle \boldsymbol{m}_h^i \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{\Pi}_h^\theta(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) + \boldsymbol{f}_h^{i+\theta}], \boldsymbol{w}_h \rangle_h$$
$$- \alpha\langle \boldsymbol{m}_h^i \times \left(\boldsymbol{m}_h^i \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^i + \theta k \boldsymbol{v}_h^i) + \boldsymbol{\Pi}_h^\theta(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}) + \boldsymbol{f}_h^{i+\theta}]\right), \boldsymbol{w}_h \rangle_h\,.$$

(ii) *Define* $\hat{\boldsymbol{m}}_h^{i+1/2} := \boldsymbol{m}_h^i + (k/2)\boldsymbol{v}_h^i$ *and compute* $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$(1+\alpha^2)\langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h \rangle_h = -\langle \boldsymbol{m}_h^{i+1/2} \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h \hat{\boldsymbol{m}}_h^{i+1/2} + \boldsymbol{\pi}_h(\hat{\boldsymbol{m}}_h^{i+1/2}) + \boldsymbol{f}_h^{i+1/2}], \boldsymbol{w}_h \rangle_h$$
$$- \alpha\langle \boldsymbol{m}_h^{i+1/2} \times \left(\hat{\boldsymbol{m}}_h^{i+1/2} \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\hat{\boldsymbol{m}}_h^{i+1/2}) + \boldsymbol{\pi}_h(\hat{\boldsymbol{m}}_h^{i+1/2}) + \boldsymbol{f}_h^{i+1/2}]\right), \boldsymbol{w}_h \rangle_h\,.$$

Output: *Sequence of discrete functions* $\left\{(\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1})\right\}_{i=0}^{J-1}$.

We note that a variant of PC2+IMEX is obtained from PC2, if also in the corrector scheme of Algorithm 1.6.5(ii) the approximation $\boldsymbol{\pi}_h(\hat{\boldsymbol{m}}_h^{i+1/2}) \approx \boldsymbol{\Pi}_h^{1/2}(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1})$ is applied. This approximation is second-order accurate, and hence, would not deteriorate the order of the scheme. Practically, PC2+IMEX is to be preferred over PC2, as besides the evaluation of $\boldsymbol{\pi}_h(\cdot)$, it requires the solution of only two sparse linear systems per time-step.

1.7 Contributions of this thesis and future work

Chapter 2

As a novel contribution, we introduce and analyze all versions of the tangent plane scheme [Alo08, AHP+14, AKST14, DPP+20] for LLG in the presence of DMI, which is the essential ingredient for the enucleation and the stabilization of magnetic skyrmions. Moreover, DMI contributions represent a challenging testing ground for numerical schemes for LLG. Indeed, besides requiring accurate adaptations in the numerical analysis, DMI determines magnetization configurations — magnetic skyrmions — that turn out to be very sensitive to small perturbations of the micromagnetic energy.

For any tangent plane algorithm, we prove that the sequence of finite element solutions, upon extraction of a subsequence, converges towards a weak solution of the problem. For the projection-free algorithm, we prove that the convergence is even unconditional, while the stability analysis requires a mild CFL-type condition on the discretization parameters and imposes the angle condition (1.25) on the underlying mesh for the projection-based approaches. The extension of the LLG analysis is not straightforward, since the DMI term involves magnetization derivatives, is neither self-adjoint nor positive definite, and requires to impose different boundary conditions on LLG, which entail a careful treatment. A by-product of our constructive analysis is the proof of existence of weak solutions, which — to the best of our knowledge — was missing in the literature for LLG with DMI. Finally, numerical experiments show that our approach can be used to study enucleation processes, stability, and dynamics of magnetic skyrmions.

Chapter 3

We extend the midpoint scheme and its analysis to more general energy contributions, in particular, covering the non-standard DMI energy contribution [Dzy58, Mor60], which is not covered by the analysis in [BP06, Bar06, Cim09, PRS18b]. In a collection of numerical experiments we accurately test the energy conservation properties of the mass-lumped midpoint scheme and extensively compare it with the tangent plane schemes from [Alo08, AKST14, HPP$^+$19, DPP$^+$20]. We conclude that the tangent plane schemes are preferable for uncritical simulations as in Section 3.4.1 or [HPP$^+$19, 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.

Besides extending the analysis of the (ideal) midpoint scheme to more general energy contributions, we provide a rigorous stability and well-posedness analysis for two fully practical versions of the original midpoint scheme, which both incorporate an efficient implicit-explicit treatment of the effective field contributions. The first practical midpoint scheme analyzed, uses a constraint-preserving fixed-point iteration in the spirit of [Bar06] to discretize the nonlinear system of the ideal midpoint scheme. The second approach is based on linearization by the Newton method. While this linearization strategy only preserves the unit-length constraint asymptotically, in a 2D numerical experiment in [BBNP14, Section 1.4.1] the authors observe a less restrictive coupling condition on the time-step size $k > 0$. Although, in our 3D numerical study we experimentally observe a coupling condition analogous to the one for the constraint preserving fixed-point iteration, namely $k = o(h^2)$, the Newton iteration is of practical interest as our study reveals that it leads to considerably lower iteration numbers. As far as theory is concerned, for convergence of the Newton iteration with canonical starting value $\boldsymbol{m}_h^i$, our analysis imposes the sufficient coupling conditions $k = o(h^{7/3})$ and $\varepsilon = \mathcal{O}(h^{3/2})$, with $\varepsilon > 0$ the nonlinear solver accuracy. Although our numerical study indicates that the first coupling condition might not be sharp and can possibly be improved to $k = o(h^2)$, our analysis provides — to the best of our knowledge — the first rigorous analysis of Newton's method applied to the midpoint scheme for three dimensional micromagnetics to the literature.

Chapter 4

In each time-step, (any variant of) the tangent plane scheme requires the solution of a linear system of equations posed in the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^i] \subsetneq \mathcal{S}^1(\mathcal{T}_h)^3$. The nature of this linear system — it is non-symmetric and posed in a time-dependent subspace $\mathcal{K}_h[\boldsymbol{m}_h^i]$ of the finite element space $\mathcal{S}^1(\mathcal{T}_h)^3$ — makes efficient numerical solution, which has not been discussed in the literature before, a non-trivial task. Finite element software packages like NGSolve [ngs] usually assemble linear systems in the full discrete space $\mathcal{S}^1(\mathcal{T}_h)^3$. Starting from such a linear system in $\mathcal{S}^1(\mathcal{T}_h)^3$, based on nodewise Householder reflections we construct the constrained linear system of the tangent plane scheme posed in $\mathcal{K}_h[\boldsymbol{m}_h^i]$, recovering the effective dimension of the problem. The tangent plane constraint aggravates the construction of suitable and effective preconditioners, which, if possible, should not

depend on the time-step, or, at least, only need an update every once in a while (after several time-steps). We derive possible preconditioners, which are (essentially) independent of the time-step, and prove linear convergence of the preconditioned GMRES algorithm. Numerical experiments underpin the theoretical findings in an academic setup, as well as in micromagnetic benchmark problems from [MUM].

Chapter 5

We improve the theoretical understanding of the predictor-corrector methods proposed in [KW18] by establishing unconditional well-posedness of both, `PC1` and `PC2`, i.e., for each time-step, the variational problems to be solved admit a unique solution, which is left open in the original paper. By closing this fundamental gap, we show that `PC1` is not only closely related to the first-order tangent plane scheme of [Alo08, BFF⁺14], but actually can even be interpreted as a slight modification of it. Our well-posedness analysis is based on a reformulation of the predictor system, which crucially exploits nodewise unit-length $\boldsymbol{m}_h^i \in \boldsymbol{\mathcal{M}}_h$. As in practice linear systems are not solved exactly, practically nodewise unit-length can formally not be satisfied for iterates of `PC2`, but only in theory. To cope with this problem, we establish a decomposition of the finite element space, which does not only allow us to prove unconditional well-posedness of the practical version of `PC2`, but also to extend the result, for both `PC1` and `PC2` (theoretical and practical), to the limit case $\alpha = 0$ (Schrödinger map equation).

Furthermore, following [BFF⁺14, PRS18b, DPP⁺20], we propose implicit-explicit versions of `PC1` and `PC2`. When considering magnetization dynamics involving the full effective field — more precisely, dynamics including the nonlocal stray field — the proposed IMEX versions `PC1+IMEX` and `PC2+IMEX` are computationally much more attractive than their original counterparts. Experimentally, we demonstrate the improved efficiency of the IMEX schemes and their applicability to a physically relevant benchmark problem. The experiments also validate their preservation of the experimental first- and second-order accuracies of `PC1` and `PC2`, respectively.

Finally, we shed some light on the stability of `PC2(+IMEX)` by a numerical study. In particular, our experiments reveal that stability of the second-order predictor-corrector method is an issue, at least if the 'degree of implicitness' $0 \leq \theta \leq 1$ is chosen far from (not chosen as) $\theta = 1/2$. From a theoretical point of view, stability and convergence of `PC2`, not addressed in [KW18], remain open also in our analysis and will be the subject of future research.

Chapter 6

As a central element of this thesis, the theoretical contributions of our research are accompanied by the development of an easy-to-use open-source software for the simulation of micromagnetic phenomena described by LLG. In Chapter 6, we present our open-source Python module *Commics* for the study of the magnetization dynamics in ferromagnetic materials via micromagnetic simulations. *Commics* implements state-of-the-art unconditionally convergent finite element methods for the numerical integration of the Landau–Lifshitz–Gilbert equation. The implementation is based on the multiphysics finite element software

Netgen/NGSolve. The simulation scripts are written in Python, which leads to very readable code and direct access to extensive post-processing. Together with documentation and example scripts, the code is freely available on GitLab [Pfe].

Future work

Since the "perfect" integrator for LLG was not proposed yet, there is demand for numerical schemes satisfying unique combinations of theoretical requirements and practical desires. Hence, the proposition of new schemes as well as the (further) improvement of existing integrators is of great interest. We point out two questions, which, despite considerable effort, remain unanswered in our analysis:

- **Stability of `PC2(+IMEX)`.** While we could not provide a rigorous stability analysis of `PC2`, our numerical study in Section 5.5.3 motivates the following future approach to the question: The experiments unambiguously reveal increased stability of simulations with `PC2`, where the parameter of implicitness $0 \leq \theta \leq 1$ is chosen close to $1/2$. Hence, in future research a possible first step towards theoretically understanding stability of `PC2` might be to establish stability of `PC2` for the special case $\theta = 1/2$, possibly under some CFL condition. This seems reasonable, as in this special case only the same highest-order term $\boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)$ appears in the predictor and the corrector of `PC2` and thus potentially cancels out (partially) in a rigorous stability analysis.

- **Coupling conditions for `MPS(+IMEX)` with Newton linearization.** Our numerical study in Section 3.4.3 confirms that in terms of iteration numbers the Newton linearization of `MPS` outperforms the linearization by a (constraint preserving) fixed-point iteration. This is in accordance with theory, predicting quadratic convergence of Newton's method for appropriate initial guesses. At the same time, in the experiments we observe that both linearization strategies are equally restrictive on the time-step size. Hence, possibly the CFL conditions for the Newton solver $k = o(h^{7/3})$ imposed in our analysis are not sharp and might be weakened to $k = o(h^2)$ in an improved analysis.

2 Convergent tangent plane integrators for the simulation of chiral magnetic skyrmion dynamics

This chapter consists of the article [HPP$^+$19] together with Gino Hrkac, Dirk Praetorius, Michele Ruggeri, Antonio Segatti and Bernhard Stiftner.

Abstract. We consider the numerical approximation of the Landau–Lifshitz–Gilbert equation, which describes the dynamics of the magnetization in ferromagnetic materials. In addition to the classical micromagnetic contributions, the energy comprises the Dzyaloshinskii–Moriya interaction, which is the most important ingredient for the enucleation and the stabilization of chiral magnetic skyrmions. We propose and analyze three tangent plane integrators, for which we prove (unconditional) convergence of the finite element solutions towards a weak solution of the problem. The analysis is constructive and also establishes existence of weak solutions. Numerical experiments demonstrate the applicability of the methods for the simulation of practically relevant problem sizes.

2.1 Introduction

2.1.1 State of the art

Magnetic skyrmions are topologically protected vortex-like magnetization configurations [NT13, FBT$^+$16, Wie16], which have been theoretically predicted [BY89, BH94, BR01, RBP06] and experimentally observed [MBJ$^+$09, RHM$^+$13] in several magnetic systems. The most important ingredient for the enucleation and the stabilization of magnetic skyrmions is the so-called Dzyaloshinskii–Moriya interaction (DMI); see [Dzy58, Mor60]. It is a short-range effect, sometimes also referred to as antisymmetric exchange, which exerts a torque on the magnetization inducing neighboring spins to be perpendicular to each other. It is thus in direct competition with the classical Heisenberg exchange interaction, which conversely favors uniform configurations. The DMI is modeled by an energy contribution, which is linear in the first spatial derivatives of the magnetization and is added to the micromagnetic energy for chiral ferromagnets. Magnetic skyrmions are currently subject of intense scientific research, which includes theoretical, computational, and experimental studies; see, e.g., [HvBM$^+$11, SCR$^+$13, KP15, HKK$^+$16, BAB$^+$17]. As for the mathematical literature, the existence of isolated skyrmions emerging as energy minimizers of two-dimensional micromagnetic models and their dynamic stability have been investigated in [Mel14, DM17], whereas chiral domain walls in ultrathin ferromagnetic films have been studied in [MS17]. The growing interest in skyrmions in the magnetic storage and magnetic

logic community is connected with their potential as possible candidate to store the bits of future devices, with the information being encoded as presence/absence of a skyrmion; see, e.g., [FCS13, TMZ+14] for the proposal of skyrmion racetrack memories, which are believed to overcome the original domain-wall-based device of [PHT08] and pave new ways in magnetic data logic [HKBB15].

A well-accepted model for the magnetization dynamics is the Landau–Lifshitz–Gilbert equation (LLG) [LL35, Gil55]. The numerical approximation of LLG poses several challenges: nonlinearities, a nonconvex pointwise constraint, an intrinsic energy law, which resembles the one of a gradient flow and combines conservative and dissipative effects, and the possible coupling with other partial differential equations (PDEs), e.g., the Maxwell equations.

The numerical integration of LLG has been the subject of several mathematical studies; see, e.g., [Pro01, KP06, GC07]. A well-established approach is represented by the integrators usually referred to as tangent plane schemes. These methods are based on equivalent reformulations of the equation in the tangent space.

The integrator proposed in [Alo08], which considers the case in which the energy only comprises the exchange contribution, requires only the solution of one linear system per time-step, is formally of first order in time, and is unconditionally convergent towards a weak solution of the problem, i.e., the numerical analysis of the scheme does not require to impose any restrictive CFL-type coupling condition on the time-step size and the spatial mesh size. The pointwise constraint is enforced by applying the nodal projection to the computed solution at each time-step. The scheme generalizes the explicit scheme proposed in [AJ06] and analyzed in [BKP08]. Implicit-explicit approaches of the algorithm of [Alo08] for the full effective field were independently introduced and analyzed in [AKT12, BFF+14]. Extensions of the scheme for the discretization of the coupling of LLG with other PDEs were studied in [LT13, BPPR14, LPPT15, BPP15]. Inspired by [Bar16], the projection-free version of the algorithm of [Alo08], which avoids the use of the nodal projection, was introduced, analyzed, and applied to the decoupled integration of the coupling of LLG with a spin diffusion equation for the spin accumulation in [AHP+14]. The violation of the constraint at the nodes of the mesh occurring in this case is uniformly controlled by the time-step size. The projection-free tangent plane scheme of [AHP+14] was combined with a FEM-BEM coupling method for the discretization of the coupling of LLG with the magnetoquasistatic Maxwell equations in full space in [FT17a]. There, assuming the existence of a unique sufficiently smooth solution, the authors proved optimal first-order convergence rates of the method. A tangent plane scheme characterized by an enhanced convergence order in time was proposed in [AKST14]. The method is unconditionally convergent and formally of (almost) second order in time. A more efficient implicit-explicit version of this method has been proposed in [DPP+20]. Adapting ideas from [BP06, Alo08], the recent work [KW18] proposes a similar predictor-corrector scheme based on a linear mass-lumped variational formulation of LLG.

2.1.2 Contributions and general outline

In this work, as a novel contribution, we introduce and analyze three tangent plane schemes for LLG in the presence of DMI. The integrators extend to this case the first-order scheme

of [Alo08] (Algorithm 2.3.1), its projection-free variant from [AHP+14] (Algorithm 2.3.2), and the (almost) second-order scheme of [AKST14] (Algorithm 2.3.3). For any algorithm, we prove that the sequence of finite element solutions, upon extraction of a subsequence, converges towards a weak solution of the problem. For the projection-free algorithm, we prove that the convergence is even unconditional, while the stability analysis requires a mild CFL-type condition on the discretization parameters and a geometric restriction on the underlying mesh for the other two approaches. The present extension of the LLG analysis is not straightforward, since the DMI term involves magnetization derivatives, is neither self-adjoint nor positive definite, and requires to impose different boundary conditions on LLG, which entail a careful treatment. A by-product of our constructive analysis is the proof of existence of weak solutions, which to our knowledge was missing in the literature. Finally, numerical experiments show that our approach can be used to study enucleation processes, stability, and dynamics of magnetic skyrmions.

The remainder of the work is organized as follows: For the convenience of the reader, we conclude this section by collecting the notation used throughout the paper. In Section 2.2, we propose an organic presentation of the physical background and the mathematical framework of the problem under consideration. In Section 2.3, we derive three tangent plane schemes and state the convergence result (Theorem 2.3.5). Section 2.4 is devoted to numerical experiments. Finally, in Section 2.5, we present the convergence analysis of the algorithms and, in particular, we establish the proof of Theorem 2.3.5.

2.1.3 Notation

We use the standard notation for Lebesgue, Sobolev, and Bochner spaces and norms; see, e.g., [Eva10, Chapter 5] or [BBF13, Chapter 2]. In the case of (spaces of) vector-valued or matrix-valued functions, we use bold letters, e.g., for any domain U, we denote both $L^2(U; \mathbb{R}^3)$ and $L^2(U; \mathbb{R}^{3\times3})$ by $\boldsymbol{L}^2(U)$. For the differential operators, we use the following notation: For a scalar function f, we denote by ∇f the gradient and by Δf the Laplace operator. For a vector-valued function $\boldsymbol{f}$, we denote by $\nabla \cdot \boldsymbol{f}$ the divergence, by $\nabla \times \boldsymbol{f}$ the curl, by $\boldsymbol{\nabla f}$ the Jacobian, and by $\boldsymbol{\Delta f}$ the vector-valued Laplace operator. Given another vector-valued function $\boldsymbol{h}$, we also define $(\boldsymbol{f} \cdot \nabla)\boldsymbol{h}$ by $[(\boldsymbol{f} \cdot \nabla)\boldsymbol{h}]_i = \boldsymbol{f} \cdot \nabla h_i$ for all $1 \leq i \leq 3$. We denote the unit sphere by $\mathbb{S}^2 = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| = 1\}$ and by $\{\mathbf{e}_i\}_{1 \leq i \leq 3} \in \mathbb{R}^3$ the standard basis of $\mathbb{R}^3$, i.e., $(\mathbf{e}_i)_j = \delta_{ij}$ for all $1 \leq i, j \leq 3$. Given a vector $\mathbf{b} \in \mathbb{R}^3$ and a matrix $\mathbf{A} \in \mathbb{R}^{3\times3}$ (with columns $\mathbf{a}_i \in \mathbb{R}^3$ for all $1 \leq i \leq 3$), we denote by $\mathbf{A} \times \mathbf{b} \in \mathbb{R}^{3\times3}$ the matrix whose columns are $\mathbf{a}_i \times \mathbf{b}$ for all $1 \leq i \leq 3$. By $C > 0$ we always denote a generic constant, which is independent of the discretization parameters, but not necessarily the same at each occurrence. We also use the notation $\lesssim$ to denote *smaller than or equal to up to a multiplicative constant*, i.e., we write $A \lesssim B$ if there exists a constant $C > 0$, which is clear from the context and always independent of the discretization parameters, such that $A \leq CB$.

2.2 Mathematical model

2.2.1 Physical background

Let $\Omega \subset \mathbb{R}^3$ be a bounded domain with boundary $\Gamma := \partial\Omega$. The dynamics of the normalized magnetization $\boldsymbol{m} = (m_1, m_2, m_3) \in \mathbb{S}^2$ is governed by LLG, which in the so-called Gilbert form reads

$$\partial_t \boldsymbol{m} = -\gamma_0 \, \boldsymbol{m} \times \boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m}) + \alpha \, \boldsymbol{m} \times \partial_t \boldsymbol{m}; \tag{2.1}$$

see [LL35, Gil55]. Here, $\gamma_0 \approx 2.21 \cdot 10^5 \, \mathrm{m/(A\,s)}$ is the rescaled gyromagnetic ratio of the electron, $0 < \alpha \leq 1$ is the dimensionless Gilbert damping parameter, and $\boldsymbol{H}_{\mathrm{eff}}$ is the energy-based effective field (in A/m), i.e., it holds that

$$\mu_0 M_{\mathrm{s}} \, \boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m}) = -\frac{\delta \mathcal{E}(\boldsymbol{m})}{\delta \boldsymbol{m}}, \tag{2.2}$$

where $\mathcal{E}(\cdot)$ is the total energy, $\mu_0 = 4\pi \cdot 10^{-7} \, \mathrm{N/A^2}$ is the vacuum permeability, and $M_{\mathrm{s}} > 0$ is the saturation magnetization (in A/m). In micromagnetics, the total energy is usually the sum of the following standard terms:

- the Heisenberg exchange contribution

$$\mathcal{E}_{\mathrm{ex}}(\boldsymbol{m}) = A \int_\Omega |\boldsymbol{\nabla}\boldsymbol{m}|^2 \mathrm{d}\boldsymbol{x},$$

where $A > 0$ denotes the exchange stiffness constant (in J/m);

- the magnetocrystalline anisotropy contribution, which for the uniaxial case reads

$$\mathcal{E}_{\mathrm{ani}}(\boldsymbol{m}) = K \int_\Omega \left[1 - (\boldsymbol{a} \cdot \boldsymbol{m})^2\right] \mathrm{d}\boldsymbol{x},$$

where $K > 0$ denotes the anisotropy constant (in J/m^3) and $\boldsymbol{a} \in \mathbb{S}^2$ is the easy axis,

- the Zeeman contribution

$$\mathcal{E}_{\mathrm{ext}}(\boldsymbol{m}) = -\mu_0 M_{\mathrm{s}} \int_\Omega \boldsymbol{H}_{\mathrm{ext}} \cdot \boldsymbol{m} \, \mathrm{d}\boldsymbol{x},$$

where $\boldsymbol{H}_{\mathrm{ext}}$ denotes an applied external field (in A/m);

- the magnetostatic contribution

$$\mathcal{E}_{\mathrm{mag}}(\boldsymbol{m}) = \frac{\mu_0}{2} \int_{\mathbb{R}^3} |\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m})|^2 \mathrm{d}\boldsymbol{x},$$

where $\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m}) = -\nabla u$ denotes the stray field (in A/m), with u being the magnetostatic potential (in A), which solves the full-space transmission problem

$$-\Delta u^{\mathrm{int}} = -M_{\mathrm{s}} \nabla \cdot \boldsymbol{m} \quad \text{in } \Omega, \tag{2.3a}$$

$$-\Delta u^{\mathrm{ext}} = 0 \quad \text{in } \mathbb{R}^3 \setminus \overline{\Omega}, \tag{2.3b}$$

$$u^{\mathrm{ext}} - u^{\mathrm{int}} = 0 \quad \text{on } \Gamma, \tag{2.3c}$$

$$(\nabla u^{\mathrm{ext}} - \nabla u^{\mathrm{int}}) \cdot \boldsymbol{n} = -M_{\mathrm{s}} \boldsymbol{m} \cdot \boldsymbol{n} \quad \text{on } \Gamma, \tag{2.3d}$$

$$u(\boldsymbol{x}) = \mathcal{O}(1/|\boldsymbol{x}|) \quad \text{as } |\boldsymbol{x}| \to \infty. \tag{2.3e}$$

Here, $\boldsymbol{n} : \Gamma \to \mathbb{S}^2$ denotes the outward-pointing unit normal vector to Γ.

In this work, the total energy $\mathcal{E}(\cdot)$ in (2.2) also comprises a term associated with the DMI [Dzy58, Mor60]. This energy contribution is phenomenologically introduced for systems with a broken symmetry due to different interface crystal configurations as a linear combination of the so-called Lifshitz invariants, i.e., the components of the chirality tensor $\mathbf{A} = (a_{ij})_{1 \leq i,j \leq 3} = \boldsymbol{\nabla m} \times \boldsymbol{m}$; see [BY89, BH94]. The choice of the appropriate DMI form depends on the crystal structure of the material and on the geometry of the sample under consideration. This choice in turn determines the specific expression of the effective field (according to (2.2)) and the boundary conditions on Γ, which are chosen in agreement with those satisfied by the solution of the Euler–Lagrange equations associated with the energy minimization problem

$$\min_{|\boldsymbol{m}|=1} \mathcal{E}(\boldsymbol{m}). \tag{2.4}$$

In micromagnetics, two main DMI forms are usually considered:

- For helimagnetic materials [Mor60], the DMI is obtained by taking as Lifshitz invariants the trace of the matrix $\mathbf{A}$, i.e.,

$$\operatorname{tr} \mathbf{A} = \sum_{1 \leq i \leq 3} a_{ii} = (\nabla \times \boldsymbol{m}) \cdot \boldsymbol{m}.$$

  The so-called *bulk* DMI energy contribution then takes the form

$$\mathcal{E}_{\mathrm{bDMI}}(\boldsymbol{m}) = D \int_{\Omega} (\nabla \times \boldsymbol{m}) \cdot \boldsymbol{m} \, \mathrm{d}\boldsymbol{x}, \tag{2.5}$$

  so that the resulting effective field term and boundary conditions on Γ are given by

$$\boldsymbol{H}_{\mathrm{eff,bDMI}}(\boldsymbol{m}) = -\frac{2D}{\mu_0 M_{\mathrm{s}}} \nabla \times \boldsymbol{m} \quad \text{and} \quad 2A \, \partial_{\boldsymbol{n}} \boldsymbol{m} + D \, \boldsymbol{m} \times \boldsymbol{n} = \boldsymbol{0}.$$

- Another type of DMI is due to the interfaces between different materials which break inversion symmetry [CL98]. For a magnetic thin film aligned with the $x_1 x_2$-plane, the Lifshitz invariants are given by

$$a_{12} - a_{21} = m_3 (\partial_1 m_1 + \partial_2 m_2) - (m_1 \, \partial_1 m_3 + m_2 \, \partial_2 m_3).$$

  The so-called *interfacial* DMI energy contribution thus takes the form

$$\mathcal{E}_{\mathrm{iDMI}}(\boldsymbol{m}) = D \int_{\Omega} [m_3 (\partial_1 m_1 + \partial_2 m_2) - (m_1 \, \partial_1 m_3 + m_2 \, \partial_2 m_3)] \, \mathrm{d}\boldsymbol{x}, \tag{2.6}$$

  so that the resulting effective field term and boundary conditions on Γ are given by

$$\boldsymbol{H}_{\mathrm{eff,iDMI}}(\boldsymbol{m}) = -\frac{2D}{\mu_0 M_{\mathrm{s}}} \begin{pmatrix} -\partial_1 m_3 \\ -\partial_2 m_3 \\ \partial_1 m_1 + \partial_2 m_2 \end{pmatrix} \quad \text{and} \quad 2A \, \partial_{\boldsymbol{n}} \boldsymbol{m} + D(\mathbf{e}_3 \times \boldsymbol{n}) \times \boldsymbol{m} = \boldsymbol{0}.$$

In (2.5)–(2.6), the constant $D \in \mathbb{R}$ is the DMI constant (in J/m^2)[1]. The sign of D determines the chirality of the system, which, in the case of a skyrmion state, defines the sense of rotation of the magnetization along the skyrmion diameter [SCR⁺13, KP15].

---

[1] Consider the energy w_D in [BH94, equation (4)] for the crystallographic class C_n ($n = 3, 4, 6$). The bulk DMI energy (2.5) corresponds to the last two terms of w_D, i.e., $D_1 = 0$ and $D_2 = D_3 = -D$ (crystallographic subclass D_n). The interfacial DMI energy (2.6) corresponds to the first term of w_D, i.e., $D_1 = D$ and $D_2 = D_3 = 0$ (crystallographic subclass C_{nv}).

2.2.2 Problem formulation

To simplify the notation, in our analysis, we restrict ourselves to the case in which the energy solely comprises exchange and DMI. We refer to [AKT12, BFF$^+$14, AHP$^+$14, AKST14, DPP$^+$20] for the design and the analysis of effective tangent plane integrators for the standard energy terms (exchange, anisotropy, Zeeman, and magnetostatic). Moreover, without loss of generality, we assume that $D > 0$ and consider the prototypical case of bulk DMI, because it is characterized by a short notation involving the curl operator. The same approach and the same results hold for all possible choices of the Lifshitz invariants.

After a suitable nondimensionalization[2], the initial boundary value problem in which we are interested takes the form

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) + \alpha\, \boldsymbol{m} \times \partial_t \boldsymbol{m} \quad \text{in } \Omega \times (0, \infty), \tag{2.7a}$$

$$2\ell_{\text{ex}}^2\, \partial_{\boldsymbol{n}} \boldsymbol{m} = -\ell_{\text{dm}}\, \boldsymbol{m} \times \boldsymbol{n} \quad \text{on } \Gamma \times (0, \infty), \tag{2.7b}$$

$$\boldsymbol{m}(0) = \boldsymbol{m}^0 \quad \text{in } \Omega, \tag{2.7c}$$

where the effective field $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m})$ is determined by the energy functional

$$\mathcal{E}(\boldsymbol{m}) = \frac{\ell_{\text{ex}}^2}{2} \|\boldsymbol{\nabla}\boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\text{dm}}}{2} \langle \nabla \times \boldsymbol{m}, \boldsymbol{m} \rangle \tag{2.8}$$

according to the relation

$$\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = -\frac{\delta\mathcal{E}(\boldsymbol{m})}{\delta\boldsymbol{m}} = \ell_{\text{ex}}^2 \boldsymbol{\Delta}\boldsymbol{m} - \ell_{\text{dm}}\nabla \times \boldsymbol{m}. \tag{2.9}$$

The positive quantities $\ell_{\text{ex}} = \sqrt{2A/(\mu_0 M_{\text{s}}^2)}$ and $\ell_{\text{dm}} = 2D/(\mu_0 M_{\text{s}}^2)$ denote the exchange length and the DMI length (both measured in m), respectively.

Since $\|\nabla \times \boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)} \leq \sqrt{2}\|\boldsymbol{\nabla}\boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)}$, using the weighted Young inequality

$$ab \leq \frac{\varepsilon a^2}{2} + \frac{b^2}{2\varepsilon} \quad \text{for any } a, b \in \mathbb{R} \text{ and } \varepsilon > 0, \tag{2.10}$$

it is easy to see that the energy (2.8) satisfies the condition

$$\frac{\ell_{\text{ex}}^2}{4}\|\boldsymbol{\nabla}\boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)}^2 - \frac{\ell_{\text{dm}}^2}{2\ell_{\text{ex}}^2}\|\boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \mathcal{E}(\boldsymbol{m}) \leq \frac{\ell_{\text{ex}}^2 + \ell_{\text{dm}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{1}{4}\|\boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)}^2. \tag{2.11}$$

For any $T > 0$, we define the space-time cylinder by $\Omega_T := \Omega \times (0, T)$. Moreover, we denote by $\langle \cdot, \cdot \rangle$ the scalar product in $\boldsymbol{L}^2(\Omega)$, by $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ the duality pairing between $\boldsymbol{H}^{-1/2}(\Gamma)$ and $\boldsymbol{H}^{1/2}(\Gamma)$, and by $\gamma_T : \boldsymbol{H}(\text{curl}, \Omega) \to \boldsymbol{H}^{-1/2}(\Gamma)$ the tangential trace operator, which satisfies $\gamma_T[\boldsymbol{u}] = \boldsymbol{u} \times \boldsymbol{n}|_\Gamma$ for any smooth function $\boldsymbol{u}$ as well as the Green formula

$$\langle\!\langle \gamma_T[\boldsymbol{u}], \boldsymbol{\phi} \rangle\!\rangle = \langle \boldsymbol{u}, \nabla \times \boldsymbol{\phi} \rangle - \langle \nabla \times \boldsymbol{u}, \boldsymbol{\phi} \rangle \quad \text{for all } \boldsymbol{u} \in \boldsymbol{H}(\text{curl}, \Omega) \text{ and } \boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega); \tag{2.12}$$

see, e.g., [BBF13, Lemma 2.1.4].

We conclude this section by extending the notion of a weak solution introduced in [AS92] to the present setting.

---

[2]We rescale the time according to the transformation $t' = \gamma_0 M_{\text{s}} t$. We define the rescaled effective field by $\boldsymbol{h}_{\text{eff}} = \boldsymbol{H}_{\text{eff}}/M_{\text{s}}$ and the rescaled energy by $\mathcal{E}' = \mathcal{E}/(\mu_0 M_{\text{s}}^2)$. However, to simplify the notation, we neglect all $'$-superscripts.

Definition 2.2.1. *Let* $\boldsymbol{m}^0 \in \boldsymbol{H}^1(\Omega)$ *satisfy* $|\boldsymbol{m}| = 1$ *a.e. in* Ω. *A vector field* $\boldsymbol{m} : \Omega \times (0,\infty) \to \mathbb{R}^3$ *is called a weak solution of* (2.7) *if, for any* $T > 0$, *the following properties are satisfied:*

(i) $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T) \cap L^\infty(0,T;\boldsymbol{H}^1(\Omega))$ *with* $|\boldsymbol{m}| = 1$ *a.e. in* Ω_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*

$$\int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t$$
$$= \ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{\nabla}\boldsymbol{m}(t), \boldsymbol{\nabla}\boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t - \ell_{\mathrm{dm}} \int_0^T \langle \nabla \times \boldsymbol{m}(t), \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t$$
$$- \frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle\!\langle \gamma_T[\boldsymbol{m}(t)], \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle\!\rangle \, \mathrm{d}t + \alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t;$$

$$(2.13)$$

(iv) It holds that

$$\mathcal{E}(\boldsymbol{m}(T)) + \alpha \int_0^T \|\partial_t \boldsymbol{m}(t)\|_{\boldsymbol{L}^2(\Omega)}^2 \mathrm{d}t \leq \mathcal{E}(\boldsymbol{m}^0). \tag{2.14}$$

The variational formulation (2.13) comes from a weak formulation of (2.7a) in the space-time domain. The boundary conditions (2.7b) are enforced as natural boundary conditions. In particular, the term with $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ arises from integrating by parts the exchange contribution and using (2.7b). The energy inequality (2.14) is a weak counterpart of the dissipative energy law

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

satisfied by any sufficiently smooth solution of (2.7).

Remark 2.2.2. *Taking the scalar product of* (2.7a) *with* $\boldsymbol{m}$, *we deduce that* $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$. *In particular, since* $\partial_t |\boldsymbol{m}|^2 = 2\,\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$, *it follows that a sufficiently smooth solution of* (2.7a) *satisfies the constraint* $|\boldsymbol{m}| = 1$, *provided that it is satisfied by the initial condition.*

2.3 Numerical algorithms and main result

In this section, we introduce three algorithms for the numerical approximation of the problem discussed in Section 2.2.2 and we state the main convergence result.

2.3.1 Preliminaries

For the time discretization, given an integer $N > 0$ and a final time $T > 0$, we consider a uniform partition of the time interval $(0,T)$ with time-step size $k := T/N$, i.e., $t_i := ik$ for all $0 \leq i \leq N$. For the spatial discretization, we assume Ω to be a polyhedral domain with Lipschitz boundary and consider a κ-quasi-uniform family $\{\mathcal{T}_h\}_{h>0}$ of regular tetrahedral

meshes of Ω parametrized by the mesh size $h > 0$, i.e., there exists $\kappa \geq 1$, independent of h, such that $\mathcal{T}_h$ is κ-shape-regular and $\kappa^{-1}h \leq \text{diam}(K)$ 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 linear polynomials on K. We consider the space

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

of piecewise linear and globally continuous functions from Ω to $\mathbb{R}$. The classical basis for this finite-dimensional linear space is given by the set of the 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 $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{N}_h$. We assume that all off-diagonal entries of the so-called stiffness matrix are nonpositive, i.e.,

$$\langle \nabla \varphi_{\boldsymbol{z}}, \nabla \varphi_{\boldsymbol{z}'} \rangle \leq 0 \quad \text{for all } \boldsymbol{z}, \boldsymbol{z}' \in \mathcal{N}_h \text{ with } \boldsymbol{z} \neq \boldsymbol{z}'. \tag{2.15}$$

This requirement is usually referred to as *angle condition*, since it is satisfied if the measure of all dihedral angles of all tetrahedra of the mesh is less than or equal to $\pi/2$.

Any solution of LLG is characterized by the nonconvex pointwise constraint $|\boldsymbol{m}| = 1$ and by the orthogonality property $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$. To mimic these properties at the discrete level, we require them to be satisfied only at the nodes of the mesh. To this end, we introduce the *set of admissible discrete magnetizations*

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

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

$$\mathcal{K}_h(\boldsymbol{\psi}_h) := \left\{ \boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3 : \boldsymbol{\psi}_h(\boldsymbol{z}) \cdot \boldsymbol{\phi}_h(\boldsymbol{z}) = 0 \text{ for all } \boldsymbol{z} \in \mathcal{N}_h \right\}, \tag{2.16}$$

which we call the *discrete tangent space* of $\boldsymbol{\psi}_h$.

2.3.2 Three tangent plane integrators

Using the well-known formula

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

(2.7a) can be formally rewritten in the form

$$\alpha \, \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - [\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) \cdot \boldsymbol{m}]\boldsymbol{m}. \tag{2.17}$$

Observing that this equation is linear with respect to the time derivative $\partial_t \boldsymbol{m}$, we introduce the free variable $\boldsymbol{v} = \partial_t \boldsymbol{m}$. For any $t \in (0, T)$, $\boldsymbol{v}(t)$ belongs to the tangent space of $\mathbb{S}^2$ at $\boldsymbol{m}(t)$. Taking this orthogonality and the expression (2.9) of the effective field into account, we obtain the following variational formulation: Find $\boldsymbol{v}(t) \in \boldsymbol{L}^2(\Omega)$ with $\boldsymbol{m}(t) \cdot \boldsymbol{v}(t) = 0$ a.e. in Ω such that

$$\begin{aligned}
\alpha \langle \boldsymbol{v}(t), \boldsymbol{\phi} \rangle &+ \langle \boldsymbol{m}(t) \times \boldsymbol{v}(t), \boldsymbol{\phi} \rangle \\
&= -\ell_{\text{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}(t), \boldsymbol{\nabla} \boldsymbol{\phi} \rangle - \frac{\ell_{\text{dm}}}{2} \langle \nabla \times \boldsymbol{m}(t), \boldsymbol{\phi} \rangle - \frac{\ell_{\text{dm}}}{2} \langle \boldsymbol{m}(t), \nabla \times \boldsymbol{\phi} \rangle
\end{aligned} \tag{2.18}$$

for all $\boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega)$ satisfying $\boldsymbol{m}(t) \cdot \boldsymbol{\phi} = 0$ a.e. in Ω. To obtain (2.18), the boundary integral which arises from integrating by parts the exchange contribution and using the boundary conditions (2.7b) is rewritten as a volume integral by using (2.12). Note that, since the test function $\boldsymbol{\phi}$ belongs to the tangent space of the sphere at $\boldsymbol{m}(t)$, in (2.18) the term corresponding to the last term (strongly nonlinear in $\boldsymbol{m}$) on the right-hand side of (2.17) vanishes.

For any time-step $0 \le i \le N - 1$, given the approximate current magnetization $\boldsymbol{m}_h^i \approx \boldsymbol{m}(t_i)$, we compute $\boldsymbol{v}_h^i \approx \boldsymbol{v}(t_i)$ by a Galerkin discretization of (2.18) based on the discrete tangent space $\mathcal{K}_h(\boldsymbol{m}_h^i)$ introduced in (2.16). The computed quantity $\boldsymbol{v}_h^i \in \mathcal{K}_h(\boldsymbol{m}_h^i)$ is then used to update the current magnetization $\boldsymbol{m}_h^i \approx \boldsymbol{m}(t_i)$ to the new value $\boldsymbol{m}_h^{i+1} \approx \boldsymbol{m}(t_{i+1})$ via a first-order time-stepping. To ensure that the discrete magnetization belongs to the set of admissible discrete magnetizations $\mathcal{M}_h$, the nodal projection is applied.

The resulting scheme, summarized in the following algorithm, extends the method proposed by [Alo08] to the present situation.

Algorithm 2.3.1 (first-order tangent plane scheme, TPS1)**.** *Input:* $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: For all $0 \le i \le N - 1$, iterate:

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

$$\alpha \langle \boldsymbol{v}_h^i, \boldsymbol{\phi}_h \rangle + \langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^i, \boldsymbol{\phi}_h \rangle + \ell_{\text{ex}}^2 \theta k \langle \boldsymbol{\nabla} \boldsymbol{v}_h^i, \boldsymbol{\nabla} \boldsymbol{\phi}_h \rangle$$
$$= -\ell_{\text{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^i, \boldsymbol{\nabla} \boldsymbol{\phi}_h \rangle - \frac{\ell_{\text{dm}}}{2} \langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{\phi}_h \rangle - \frac{\ell_{\text{dm}}}{2} \langle \boldsymbol{m}_h^i, \nabla \times \boldsymbol{\phi}_h \rangle. \tag{2.19}$$

(ii) *Define $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$ by $\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \dfrac{\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})}{|\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})|}$ for all $\boldsymbol{z} \in \mathcal{N}_h$.*

Output: Sequence of discrete functions $\left\{ (\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1}) \right\}_{0 \le i \le N-1}$.

In (2.19), the parameter $0 \le \theta \le 1$ modulates the 'degree of implicitness' of the method in the treatment of the leading-order exchange contribution of the effective field.

In the following algorithm, we state a projection-free variant of Algorithm 2.3.1, where step (ii) is replaced by a simple linear first-order time-stepping. Note that, omitting the nodal projection, the pointwise constraint $|\boldsymbol{m}| = 1$ is not explicitly enforced by the numerical scheme.

Algorithm 2.3.2 (projection-free first-order tangent plane scheme, PF-TPS1)**.** *Input:* $\boldsymbol{m}_h^0 \in \mathcal{S}^1(\mathcal{T}_h)^3$.
Loop: For all $0 \le i \le N - 1$, iterate:

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

$$\alpha \langle \boldsymbol{v}_h^i, \boldsymbol{\phi}_h \rangle + \langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^i, \boldsymbol{\phi}_h \rangle + \ell_{\text{ex}}^2 \theta k \langle \boldsymbol{\nabla} \boldsymbol{v}_h^i, \boldsymbol{\nabla} \boldsymbol{\phi}_h \rangle$$
$$= -\ell_{\text{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^i, \boldsymbol{\nabla} \boldsymbol{\phi}_h \rangle - \frac{\ell_{\text{dm}}}{2} \langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{\phi}_h \rangle - \frac{\ell_{\text{dm}}}{2} \langle \boldsymbol{m}_h^i, \nabla \times \boldsymbol{\phi}_h \rangle. \tag{2.20}$$

(ii) *Define $\boldsymbol{m}_h^{i+1} := \boldsymbol{m}_h^i + k\boldsymbol{v}_h^i \in \mathcal{S}^1(\mathcal{T}_h)^3$.*

Output: Sequence of discrete functions $\left\{(\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1})\right\}_{0 \le i \le N-1}.$

The idea of removing the nodal projection from the tangent plane scheme goes back to [AHP$^+$14] for LLG and has been inspired by [Bar16], where the same principle is applied to a certain class of geometrically constrained PDEs, e.g., the harmonic map heat flow.

In [AKST14], the authors extend the tangent plane scheme of [Alo08] to improve the formal convergence order in time of the method. If the tangential update $\boldsymbol{v}$ takes the form

$$\boldsymbol{v}(t) = \partial_t \boldsymbol{m}(t) + \frac{k}{2} \mathbb{P}_{\boldsymbol{m}(t)}[\partial_{tt} \boldsymbol{m}(t)], \tag{2.21}$$

where $\mathbb{P}_{\boldsymbol{u}} : \mathbb{R}^3 \to \operatorname{span}(\boldsymbol{u})^\perp$ denotes the orthogonal projection onto $\operatorname{span}(\boldsymbol{u})^\perp$ for any $\boldsymbol{u} \in \mathbb{S}^2$, a sufficiently smooth solution $\boldsymbol{m}$ of LLG satisfies the Taylor expansion

$$\boldsymbol{m}(t+k) = \frac{\boldsymbol{m}(t) + k\boldsymbol{v}(t)}{|\boldsymbol{m}(t) + k\boldsymbol{v}(t)|} + \mathcal{O}(k^3). \tag{2.22}$$

Differentiating (2.17) with respect to time and proceeding as in [AKST14, Section 6], one obtains that, up to a residual term of order $\mathcal{O}(k^2)$, the tangential update (2.21) can be characterized as the solution of the following variational problem: Find $\boldsymbol{v}(t) \in \boldsymbol{H}^1(\Omega)$ with $\boldsymbol{m}(t) \cdot \boldsymbol{v}(t) = 0$ a.e. in Ω such that

$$\alpha\langle \boldsymbol{v}(t), \boldsymbol{\phi}\rangle + \frac{1}{2}k\langle[-\ell_{\mathrm{ex}}^2|\boldsymbol{\nabla}\boldsymbol{m}(t)|^2 - \ell_{\mathrm{dm}}(\nabla \times \boldsymbol{m}(t)) \cdot \boldsymbol{m}(t)]\boldsymbol{v}(t), \boldsymbol{\phi}\rangle$$

$$+ \langle \boldsymbol{m}(t) \times \boldsymbol{v}(t), \boldsymbol{\phi}\rangle + \frac{\ell_{\mathrm{ex}}^2}{2}k\langle\boldsymbol{\nabla}\boldsymbol{v}(t), \boldsymbol{\nabla}\boldsymbol{\phi}\rangle + \frac{\ell_{\mathrm{dm}}}{4}k\langle\nabla \times \boldsymbol{v}(t), \boldsymbol{\phi}\rangle + \frac{\ell_{\mathrm{dm}}}{4}k\langle\boldsymbol{v}(t), \nabla \times \boldsymbol{\phi}\rangle$$

$$= -\ell_{\mathrm{ex}}^2\langle\boldsymbol{\nabla}\boldsymbol{m}(t), \boldsymbol{\nabla}\boldsymbol{\phi}\rangle - \frac{\ell_{\mathrm{dm}}}{2}\langle\nabla \times \boldsymbol{m}(t), \boldsymbol{\phi}\rangle - \frac{\ell_{\mathrm{dm}}}{2}\langle\boldsymbol{m}(t), \nabla \times \boldsymbol{\phi}\rangle$$

$$\tag{2.23}$$

for all $\boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega)$ satisfying $\boldsymbol{m}(t) \cdot \boldsymbol{\phi} = 0$ a.e. in Ω. To obtain an effective numerical method, we use the same predictor-corrector approach used for Algorithms 2.3.1–2.3.2: For any time-step $0 \le i \le N-1$, given the approximation $\boldsymbol{m}_h^i \approx \boldsymbol{m}(t_i)$, we compute $\boldsymbol{v}_h^i \approx \boldsymbol{v}(t_i)$ by a Galerkin discretization of (2.23) based on $\mathcal{K}_h(\boldsymbol{m}_h^i)$. Then, with (2.22) in mind, we define $\boldsymbol{m}_h^{i+1} \approx \boldsymbol{m}(t_{i+1})$ in $\mathcal{M}_h$ as the nodal projection of $\boldsymbol{m}_h^i + k\boldsymbol{v}_h^i$.

However, in order to obtain a well-defined scheme, following [AKST14], we perform two higher-order modifications of (2.23). Firstly, to ensure the well-posedness of the variational problem, we proceed as follows: Given $M > 0$, we define the cut-off function $W_M : \mathbb{R} \to \mathbb{R}$ by

$$W_M(s) = \begin{cases} \alpha + k\min\{s, M\}/2 & \text{if } s \ge 0, \\ 2\alpha^2/(2\alpha + k\min\{-s, M\}) & \text{if } s < 0. \end{cases}$$

By construction, it holds that

$$W_M(s) \ge 2\alpha^2/(2\alpha + Mk) \quad \text{and} \quad \left|W_M(s) - \alpha\right| \le Mk/2 \quad \text{for all } s \in \mathbb{R}; \tag{2.24}$$

see, e.g., [DPP$^+$20, Lemma 12]. In the variational formulation (2.23), we then replace

$$\alpha\langle\boldsymbol{v}(t), \boldsymbol{\phi}\rangle + \frac{1}{2}k\langle[-\ell_{\mathrm{ex}}^2|\boldsymbol{\nabla}\boldsymbol{m}(t)|^2 - \ell_{\mathrm{dm}}(\nabla \times \boldsymbol{m}(t)) \cdot \boldsymbol{m}(t)]\boldsymbol{v}(t), \boldsymbol{\phi}\rangle \text{ by } \langle W_M(\lambda(\boldsymbol{m}(t)))\boldsymbol{v}(t), \boldsymbol{\phi}\rangle,$$

where

$$\lambda(\boldsymbol{m}) = \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) \cdot \boldsymbol{m} = -\ell_{\text{ex}}^2 |\boldsymbol{\nabla m}|^2 - \ell_{\text{dm}}(\nabla \times \boldsymbol{m}) \cdot \boldsymbol{m}.$$

Note that the function $\lambda(\boldsymbol{m})$ is also the Lagrange multiplier associated with the constraint $|\boldsymbol{m}| = 1$ in the constrained minimization problem (2.4). If $M > 0$ is sufficiently large, this modification introduces a consistency error of order $\mathcal{O}(k^2)$ in (2.23). In particular, to ensure this, we define $M : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ by $M(k) := |k \log k|^{-1}$ for all $k > 0$. Note that M satisfies the convergence properties

$$M(k) \to \infty \quad \text{and} \quad M(k)k \to 0 \quad \text{as } k \to 0. \tag{2.25}$$

Secondly, in the variational formulation (2.23), we replace

$$\frac{\ell_{\text{ex}}^2}{2} k \langle \boldsymbol{\nabla v}(t), \boldsymbol{\nabla \phi} \rangle \quad \text{by} \quad \frac{\ell_{\text{ex}}^2}{2} [1 + \rho(k)] k \langle \boldsymbol{\nabla v}(t), \boldsymbol{\nabla \phi} \rangle,$$

where the stabilization function $\rho : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is defined by $\rho(k) := |k \log k|$ for all $k > 0$. This artificial stabilization introduces a formal consistency error of order $\mathcal{O}(k^{2-\varepsilon})$ for any $0 < \varepsilon < 1$.

In the following algorithm, we summarize the proposed extension of the tangent plane scheme of [AKST14] to the present setting.

Algorithm 2.3.3 ((almost) second-order tangent plane scheme, TPS2). *Input:* $\boldsymbol{m}_h^0 \in \mathcal{M}_h$. *Loop: For all* $0 \le i \le N - 1$, *iterate:*

(i) *Set* $\lambda_h^i = -\ell_{\text{ex}}^2 |\boldsymbol{\nabla m}_h^i|^2 - \ell_{\text{dm}}(\nabla \times \boldsymbol{m}_h^i) \cdot \boldsymbol{m}_h^i$.

(ii) *Compute* $\boldsymbol{v}_h^i \in \mathcal{K}_h(\boldsymbol{m}_h^i)$ *such that, for all* $\boldsymbol{\phi}_h \in \mathcal{K}_h(\boldsymbol{m}_h^i)$, *it holds that*

$$\langle W_{M(k)}(\lambda_h^i)\boldsymbol{v}_h^i, \boldsymbol{\phi}_h \rangle + \langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^i, \boldsymbol{\phi}_h \rangle + \frac{\ell_{\text{ex}}^2}{2} k[1 + \rho(k)]\langle \boldsymbol{\nabla v}_h^i, \boldsymbol{\nabla \phi}_h \rangle$$

$$+ \frac{\ell_{\text{dm}}}{4} k \langle \boldsymbol{v}_h^i, \nabla \times \boldsymbol{\phi}_h \rangle + \frac{\ell_{\text{dm}}}{4} k \langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{\phi}_h \rangle \tag{2.26}$$

$$= -\ell_{\text{ex}}^2 \langle \boldsymbol{\nabla m}_h^i, \boldsymbol{\nabla \phi}_h \rangle - \frac{\ell_{\text{dm}}}{2} \langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{\phi}_h \rangle - \frac{\ell_{\text{dm}}}{2} \langle \boldsymbol{m}_h^i, \nabla \times \boldsymbol{\phi}_h \rangle.$$

(iii) *Define* $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$ *by* $\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) := \dfrac{\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})}{|\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})|}$ *for all* $\boldsymbol{z} \in \mathcal{N}_h$.

Output: Sequence of discrete functions $\big\{(\boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1})\big\}_{0 \le i \le N-1}$.

For the proof that the three proposed algorithms are well-posed (if the time-step size is sufficiently small in the case of TPS2), we refer to Proposition 2.5.1 below.

Remark 2.3.4. *The natural starting point for a hypothetical projection-free version of TPS2 would be the expansion*

$$\boldsymbol{m}(t + k) = \boldsymbol{m}(t) + k\boldsymbol{v}(t) + \mathcal{O}(k^3),$$

for which a nontangential *update of the form*

$$\boldsymbol{v}(t) = \partial_t \boldsymbol{m}(t) + \frac{k}{2} \partial_{tt} \boldsymbol{m}(t)$$

would be required. In particular, it is not clear how to apply the tangent plane paradigm to this situation, where the update $\boldsymbol{v}(t)$ has a nonzero component parallel to $\boldsymbol{m}(t)$ in general, i.e., $\boldsymbol{m}(t) \cdot \boldsymbol{v}(t) = \frac{k}{2}\partial_{tt}\boldsymbol{m}(t) \cdot \boldsymbol{m}(t) \neq 0$, which needs to be taken into account in order to achieve a second-order accuracy.

2.3.3 Convergence result

From any algorithm, we obtain two sequences of discrete functions $\{\boldsymbol{m}_h^i\}_{0 \leq i \leq N}$ as well as $\{\boldsymbol{v}_h^i\}_{0 \leq i \leq N-1}$. We define the piecewise linear time reconstruction $\boldsymbol{m}_{hk}$ and the piecewise constant time reconstructions $\boldsymbol{m}_{hk}^{\pm}$ and $\boldsymbol{v}_{hk}^-$ by

$$
\begin{aligned}
\boldsymbol{m}_{hk}(t) &:= \frac{t - t_i}{k}\boldsymbol{m}_h^{i+1} + \frac{t_{i+1} - t}{k}\boldsymbol{m}_h^i, \\
\boldsymbol{m}_{hk}^-(t) &:= \boldsymbol{m}_h^i, \quad \boldsymbol{m}_{hk}^+(t) := \boldsymbol{m}_h^{i+1}, \quad \text{and} \quad \boldsymbol{v}_{hk}^-(t) := \boldsymbol{v}_h^i
\end{aligned}
\tag{2.27}
$$

for all $0 \leq i \leq N - 1$ and $t \in [t_i, t_{i+1})$. The following theorem, which is the main result of the paper, states that the time reconstructions $\boldsymbol{m}_{hk}$ obtained by the three algorithms converge in an appropriate sense towards a weak solution of (2.7) as $h, k \to 0$.

Theorem 2.3.5. *Let the approximate initial condition satisfy the convergence property*

$$
\boldsymbol{m}_h^0 \to \boldsymbol{m}^0 \quad \text{in } \boldsymbol{H}^1(\Omega) \quad \text{as } h \to 0.
\tag{2.28}
$$

Moreover, for each algorithm, consider the following specific assumptions:

- *For TPS1 (Algorithm 2.3.1), assume that the angle condition (2.15) is satisfied, that $1/2 \leq \theta \leq 1$, and that it holds that $k/h \to 0$ as $h, k \to 0$.*

- *For PF-TPS1 (Algorithm 2.3.2), assume that $1/2 < \theta \leq 1$.*

- *For TPS2 (Algorithm 2.3.3), assume that the angle condition (2.15) is satisfied and that it holds that $k/h \to 0$ as $h, k \to 0$.*

Then, for each algorithm, there exist a weak solution $\boldsymbol{m}$ of (2.7) and a subsequence of $\{\boldsymbol{m}_{hk}\}$ which converges weakly in $\boldsymbol{H}^1(\Omega_T)$ towards $\boldsymbol{m}$ as $h, k \to 0$.

Remark 2.3.6. *(i) For the sake of brevity, we have considered the case in which the energy consists of only exchange and DMI. Adopting the implicit-explicit approaches of [AKT12, BFF+14, AHP+14, DPP+20], the schemes and the convergence result of Theorem 2.3.5 can be extended to the case in which also the standard lower-order energy terms (and their discretizations) are included into the setting.*

(ii) One important aspect of the research on numerical integrators for LLG is related to the development of unconditionally convergent methods, for which the numerical analysis does not require to impose any CFL-type condition on h and k. Theorem 2.3.5 states that this goal is achieved by PF-TPS1. For TPS1 and TPS2, our analysis requires a mild CFL condition, which arises from the use of the nodal projection and the presence of the DMI. If the energy comprises only the standard micromagnetic contributions (exchange, uniaxial anisotropy, Zeeman, and magnetostatic), but no DMI, then the convergence towards

a weak solution of LLG is unconditional also for TPS1 (for $1/2 < \theta \leq 1$) and TPS2; see [AKT12, BFF$^+$14, AKST14, DPP$^+$20].

(iii) Since the treatment of the DMI requires the CFL condition $k/h \to 0$ as $h, k \to 0$ in our analysis, the result of Theorem 2.3.5 holds for TPS2 also without artificial stabilization, i.e., $\rho \equiv 0$; see [AKST14, DPP$^+$20] for more details. In this case, TPS2 is of full second order in time.

(iv) In Theorem 2.3.5, we state the best result that we are able to prove in terms of stability, i.e., with the weakest CFL condition on the discretization parameters. The same convergence result can be also established at the price of more severe restrictions. In particular, the result of Theorem 2.3.5 holds

(a) *without angle condition (2.15) for TPS1 and TPS2, if $k/h^2 \to 0$ as $h, k \to 0$;*

(b) *also for $0 \leq \theta < 1/2$ for TPS1 and PF-TPS1, if $k/h^2 \to 0$ as $h, k \to 0$;*

(c) *also for $\theta = 1/2$ for PF-TPS1, if $k/h \to 0$ as $h, k \to 0$.*

2.4 Numerical experiments

Before proceeding with the convergence analysis, we aim to show the effectivity of the proposed algorithms with three numerical experiments. The computations presented in this section have been performed with our micromagnetic software Commics [Pfe, PRS$^+$20]. Our Python code is based on the open-source finite element library Netgen/NGSolve [ngs]. The computation of the stray field, i.e., the numerical solution of the transmission problem (2.3), is based on the hybrid FEM-BEM method of [FK90], which requires the evaluation of the double-layer integral operator associated with the Laplace equation; see, e.g., [BFF$^+$14, Section 4.4.1] or [PRS18b, Algorithm 12]. This part of the code exploits the open-source Galerkin boundary element library BEM++ [ŚBA$^+$15]. For all three schemes, to discretize the classical lower-order contributions (anisotropy, Zeeman, and magnetostatic), we follow the explicit approaches of [BFF$^+$14, AHP$^+$14, DPP$^+$20]. Magnetization configurations are visualized with ParaView [AGL05].

2.4.1 Comparison of the integrators

We discretize the rescaled form (2.7) of LLG for a rectangular cuboid Ω of dimensions $80\,\mathrm{nm} \times 80\,\mathrm{nm} \times 10\,\mathrm{nm}$, material parameters $\ell_{\mathrm{ex}} = 10\,\mathrm{nm}$, $\ell_{\mathrm{dm}} = 20\,\mathrm{nm}$, and $\alpha = 0.08$, dimensionless final time $T = 200$, and constant initial condition $\boldsymbol{m}^0 \equiv (q, -q, \sqrt{1 - 2q^2})$ with $q = 0.01$. For snapshots of the resulting magnetization dynamics, we refer to Figure 2.1.

For the spatial discretization, we consider three types of tetrahedral meshes; see Figure 2.2. For meshes of type I, the domain is first uniformly decomposed into cubes. Then, each cube is split into six tetrahedra in such a way that any tetrahedron has three mutually perpendicular edges. Any mesh of this type satisfies the angle condition (2.15); see [Bar05, Lemma 3.5]. Meshes of type II are obtained from the previous one bisecting the longest edge of each of the six tetrahedra i.e., the main diagonal of the original cube. As a result the cube is uniformly split into twelve tetrahedra. Meshes of this type do not satisfy (2.15). For

$$t = 0 \qquad t = 15 \qquad t = 27 \qquad t = 38 \qquad t = 51 \qquad t = 100 \qquad t = 150 \qquad t = 200$$
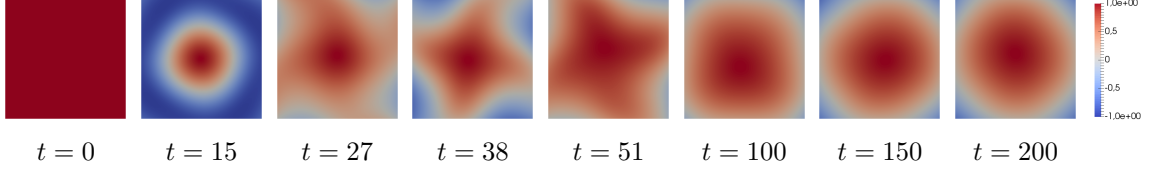
Figure 2.1: Experiment of Section 2.4.1. Snapshots of the magnetization dynamics. The color scale refers to the third component m_3 of the magnetization.

$\qquad\qquad$ (a) $\qquad\qquad\qquad\qquad$ (b) $\qquad\qquad\qquad\qquad$ (c)
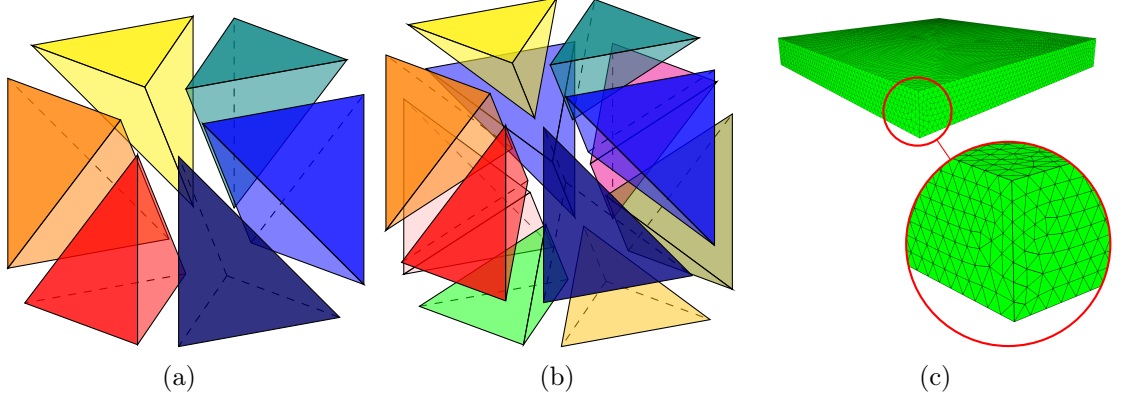
Figure 2.2: Experiment of Section 2.4.1. Mesh types: (a) Type I; (b) Type II; (c) Type III.

type III, we consider unstructured meshes obtained with Netgen, which are generated with the advancing front method (see [Sch97] for details) and in general do not satisfy (2.15).

In Figure 2.3, we plot the time evolutions of the third component of the spatially averaged magnetization of the sample, i.e., $\langle m_3(t) \rangle = |\Omega|^{-1} \int_\Omega m_3(\boldsymbol{x}, t) \, \mathrm{d}\boldsymbol{x}$, and the energy (2.8) obtained with the three algorithms for different mesh types and sizes, and a constant time-step size $k = \gamma_0 10^{-7} \approx 0.0221$. In Figure 2.3a–2.3b, we compare the results obtained with TPS1 ($\theta = 1$), PF-TPS1 ($\theta = 1$), and TPS2 for the mesh types I–III. For each mesh type, we consider a mesh size of $h \approx 3.46 \, \mathrm{nm}$. Note that the meshes of types II–III violate (2.15). In Figure 2.3c, we compare the results obtained with TPS1 ($\theta = 1/2$) and TPS2. We consider a structured mesh of type I and compare the results obtained for different mesh sizes.

Although the convergence result of Theorem 2.3.5 does not cover meshes of types II–III for TPS1 and TPS2, the numerical results show that, in terms of stability, the methods behave identically, independently of the mesh type used. To better understand this aspect, we also monitored *a posteriori* the validity of the inequality

$$\|\boldsymbol{\nabla} \boldsymbol{m}_h^{i+1}\|_{\boldsymbol{L}^2(\Omega)} \le \|\boldsymbol{\nabla}(\boldsymbol{m}_h^i + k\boldsymbol{v}_h^i)\|_{\boldsymbol{L}^2(\Omega)},$$

which is the inequality effectively used in the stability analysis of TPS1 and TPS2; see Proposition 2.5.2 and Proposition 2.5.4 below, respectively. It turned out that the inequality is always satisfied, even for meshes violating (2.15).

The omission of the nodal projection in PF-TPS1 manifests itself as a phase error in the evolution of $\langle m_3 \rangle$ accumulating over time (see Figure 2.3a), and as a lower energy level of
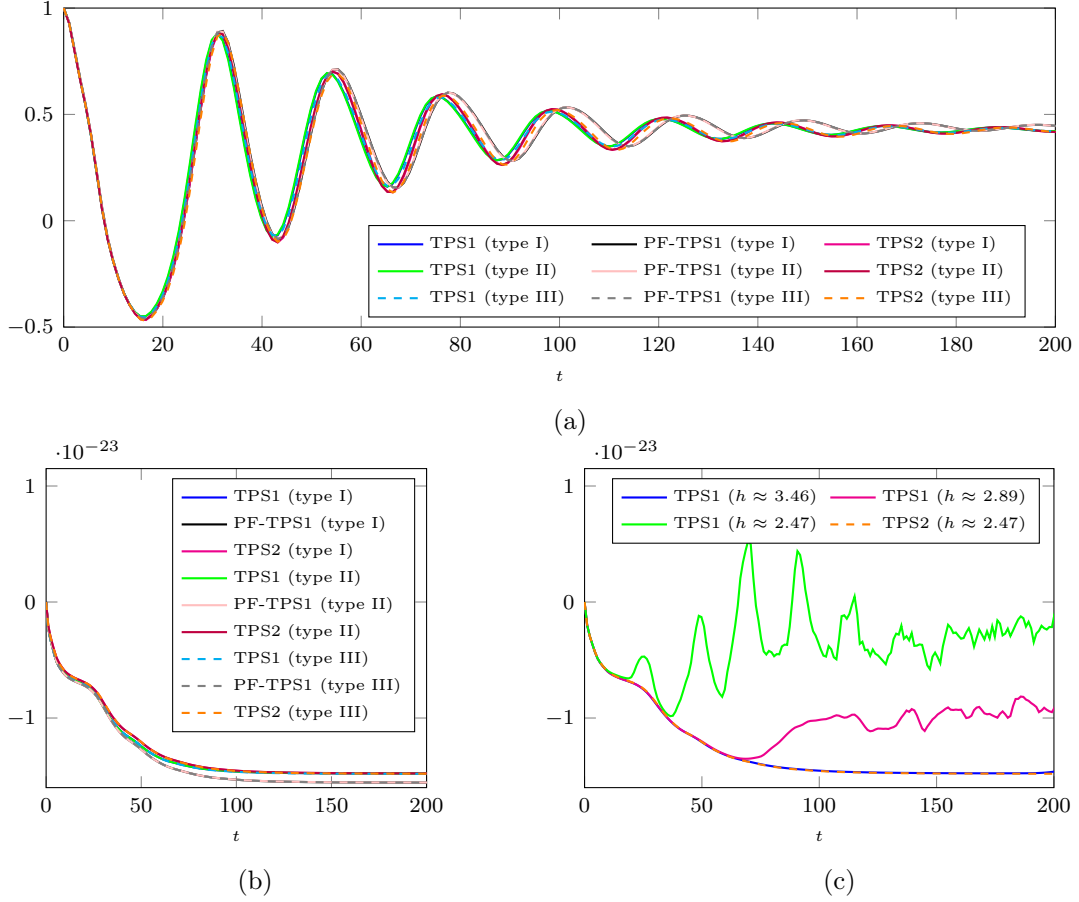
Figure 2.3: Experiment of Section 2.4.1. Time evolutions of $\langle m_3 \rangle$ and $\mathcal{E}(\boldsymbol{m})$: (a) Time evolution of $\langle m_3 \rangle$ for all schemes, all mesh types, and a fixed mesh size of $h \approx 3.46\,\text{nm}$; (b) Time evolution of $\mathcal{E}(\boldsymbol{m})$ for all schemes, all mesh types; and a fixed mesh size of $h \approx 3.46\,\text{nm}$; (c) Time evolution of $\mathcal{E}(\boldsymbol{m})$ for TPS1 ($\theta = 1/2$) and TPS2, a mesh of type I, and different mesh sizes.

the final magnetization configuration (see Figure 2.3b). However, the overall qualitative outcome of the experiment is preserved.

The results also show that TPS1 with $\theta = 1$ is more dissipative than TPS2. However, the choice of $\theta = 1/2$, which would be favorable from an energetic point of view (no artificial damping), is not feasible, because it affects the stability of the scheme; see Figure 2.3b. The instability is more severe for smaller mesh sizes, giving numerical evidence of the CFL-condition required for stability in this case; see Remark 2.3.6(ii).

Finally, in Figure 2.4, we study the violation of the unit-length constraint which occurs for PF-TPS1. We consider a structured mesh of type I with mesh size $h \approx 4.33\,\text{nm}$ and plot the error $\left\| \mathcal{I}_h \left[|\boldsymbol{m}_{hk}^+(T)|^2 \right] - 1 \right\|_{L^1(\Omega)}$ for different time-step sizes k. Note that this error is identically zero for TPS1 and TPS2, because of the nodal projection. We observe a linear dependence of the error on k which is in total agreement with the theory, see estimate (2.50)
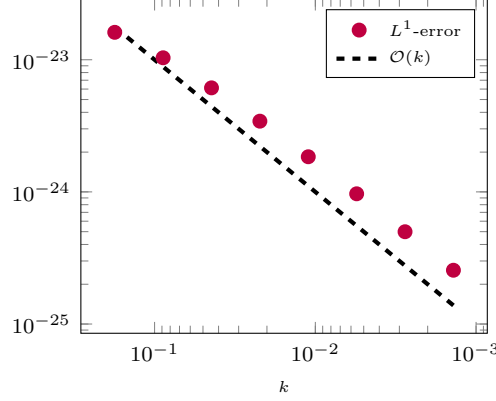
Figure 2.4: Experiment of Section 2.4.1. Empirical convergence rate as $k \to 0$ for the error $\big\|\mathcal{I}_h\big[|\boldsymbol{m}_{hk}^+(T)|^2\big] - 1\big\|_{L^1(\Omega)}$ for PF-TPS1 ($\theta = 1$).

in the proof of Proposition 2.5.6 below.

For numerical experiments testing the experimental convergence rates in time of the schemes (in the absence of DMI), we refer to [Rug16, Section 6.2.1] (for TPS1 and PF-TPS1) and [DPP+20, Section 7.1] (for TPS1 and TPS2). There, the observed rates with respect to a reference solution match the formal consistency error of the schemes, i.e., first-order convergence for TPS1 and PF-TPS1, second-order convergence for TPS2. A similar numerical study for the present model problem (which includes DMI) confirms the first-order convergence for TPS1 and PF-TPS1, but does not reveal a full second-order convergence for TPS2. We believe that this is due to a lack of regularity of the solution in time.

2.4.2 Stability of isolated skyrmions in nanodisks

We reproduce a numerical experiment from [SCR+13]. We investigate the relaxed states of a thin nanodisk of diameter $80\,\mathrm{nm}$ (aligned with $x_1 x_2$-plane) and thickness $0.4\,\mathrm{nm}$ (x_3-direction) centered at $(0,0,0)$ for different values of the DMI constant and initial conditions. The effective field in (2.1) consists of exchange interaction, perpendicular uniaxial anisotropy, interfacial DMI, and stray field, i.e.,

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

The values of the involved material parameters mimic those of cobalt: $M_{\mathrm{s}} = 5.8 \cdot 10^5\,\mathrm{A/m}$, $\alpha = 0.3$, $A = 1.5 \cdot 10^{-11}\,\mathrm{J/m}$, $K = 8 \cdot 10^5\,\mathrm{J/m^3}$, and $\boldsymbol{a} = (0,0,1)$. For the DMI constant, we consider the values $D = 0, 1, \ldots, 8\,\mathrm{mJ/m^2}$. We test two different initial magnetization configurations:

(i) a uniform out-of-plane ferromagnetic state, i.e., $\boldsymbol{m}^0 \equiv (0,0,1)$,

(ii) 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 $r \in [0, 15]\,\mathrm{nm}$ and $\boldsymbol{m}^0(\boldsymbol{x}) = (0,0,1)$ if $r \in (15, 40]\,\mathrm{nm}$.

For all simulations, we choose $T = 2\,\mathrm{ns}$ for $D = 0, \ldots, 6\,\mathrm{mJ/m^2}$ and $T = 5\,\mathrm{ns}$ for $D = 7$, 8 $\mathrm{mJ/m^2}$, which experimentally turn out to be sufficiently large times to relax the system. The computational domain is discretized by a regular partition consisting of $32\,575$ tetrahedra (mesh size of $1\,\mathrm{nm}$). For the time discretization, we consider a uniform partition of the time interval $(0, T)$ with a time-step size of $0.1\,\mathrm{ps}$.
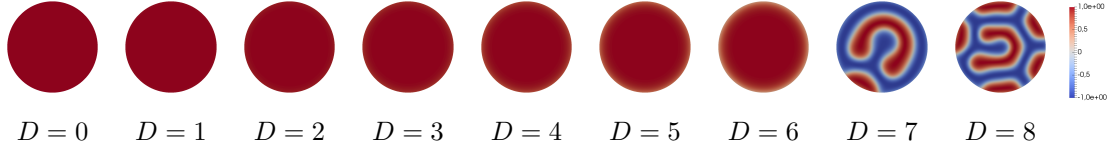
Figure 2.5: Experiment of Section 2.4.2. Magnetization m_3 of the relaxed state for the uniform out-of-plane initial condition (i) and different values of the DMI constant (in $\mathrm{mJ/m^2}$). The pictures refer to the states computed with TPS1 ($\theta = 1$).
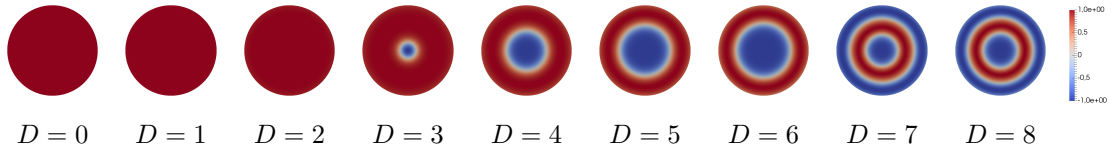
Figure 2.6: Experiment of Section 2.4.2. Magnetization m_3 of the relaxed state for the skyrmion-like initial condition (ii) and different values of the DMI constant (in $\mathrm{mJ/m^2}$). The pictures refer to the states computed with TPS1 ($\theta = 1$).

In the case of the uniform out-of-plane initial condition, the stable state remains a quasi-uniform ferromagnetic state for the values $D = 0, \ldots, 6\,\mathrm{mJ/m^2}$ and turns into a multidomain state for the values $D = 7$, $8\,\mathrm{mJ/m^2}$; see Figure 2.5. For $D = 0, \ldots, 6\,\mathrm{mJ/m^2}$, the slight decrease of the total energy for increasing values of D corresponds to an inward tilt of the magnetization on the boundary of the disk. In the case of the skyrmion-like initial condition, the stable state is a quasi-uniform ferromagnetic state for the values $D = 0, 1$, $2\,\mathrm{mJ/m^2}$, a skyrmion for the values $D = 3, \ldots, 6\,\mathrm{mJ/m^2}$, and a multidomain state for the values $D = 7$, $8\,\mathrm{mJ/m^2}$; see Figure 2.6. The skyrmion size, i.e., the diameter of the circle $\{m_3 = 0\}$ in the $x_1 x_2$-plane, increases from the minimum value of circa $14\,\mathrm{nm}$ for $D = 3\,\mathrm{mJ/m^2}$ to the maximum value of circa $48\,\mathrm{nm}$ for $D = 6\,\mathrm{mJ/m^2}$. As observed in [SCR+13], the fact that for $D = 3, \ldots, 6\,\mathrm{mJ/m^2}$, which are realistic values for the DMI constant, both the ferromagnetic state and the skyrmion state can be stabilized is very relevant for applications. Indeed, this bistability can be exploited to code the information in future recording devices (the presence and the absence of a skyrmion can be used to encode one bit); see, e.g., [FCS13, TMZ+14].

In Figure 2.7, we plot the total energy of the relaxed state for different values of the DMI constant. The energy values obtained with TPS1 (the results refer to the case $\theta = 1$) and TPS2 are in perfect quantitative agreement with each other and with those reported in [SCR+13, Figure 1]. The use of PF-TPS1 preserves the qualitative outcome of the experiment, but the quantitative agreement of the energy values with those of [SCR+13,
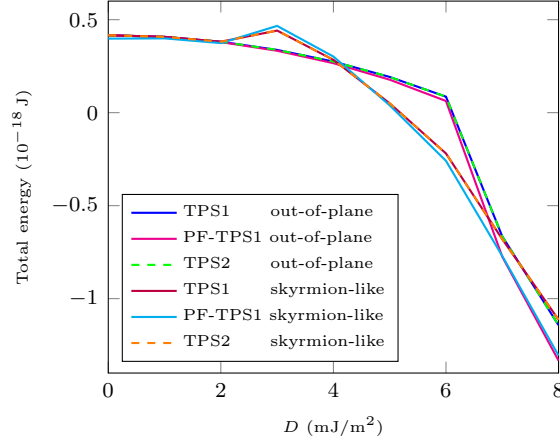
Figure 2.7: Experiment of Section 2.4.2. Total energy of the relaxed state as a function of the DMI constant D for the two considered initial conditions and the three proposed algorithms.

Figure 1], as a result of the violation of the pointwise constraint $|\boldsymbol{m}| = 1$, is inevitably lost.

2.4.3 Field-induced dynamics of skyrmions in nanodisks

We numerically investigate the stability and the induced dynamics of isolated magnetic skyrmions in helimagnetic materials in response to an applied field pulse. The sample under consideration is a magnetic nanodisk of diameter 140 nm ($x_1 x_2$-plane) and thickness 10 nm (x_3-direction). The effective field in (2.1) consists of exchange interaction, bulk DMI, applied external field, and stray field, i.e.,

$$\boldsymbol{H}_{\text{eff}}(\boldsymbol{m}) = \frac{2A}{\mu_0 M_{\text{s}}} \boldsymbol{\Delta} \boldsymbol{m} - \frac{2D}{\mu_0 M_{\text{s}}} \nabla \times \boldsymbol{m} + \boldsymbol{H}_{\text{ext}} + \boldsymbol{H}_{\text{s}}(\boldsymbol{m}).$$

We use the material parameters of iron-germanium (FeGe), i.e., $A = 8.78 \cdot 10^{-12}$ J/m, $D = 1.58 \cdot 10^{-3}$ J/m², and $M_{\text{s}} = 3.84 \cdot 10^5$ A/m; see, e.g., [BAB$^+$17]. The initial condition for our experiment is obtained by setting $\boldsymbol{H}_{\text{ext}} \equiv (0, 0, 0)$ and relaxing a uniform out-of-plane ferromagnetic state $\boldsymbol{m}^0 \equiv (0, 0, 1)$ for 3 ns. For the relaxation process, we choose the large value $\alpha = 1$ for the Gilbert damping constant, since we are not interested in the precise magnetization dynamics. The resulting relaxed state is the skyrmion depicted in Figure 2.8a. Starting from this configuration, we perturb the system from its equilibrium by applying an in-plane field pulse $\boldsymbol{H}_{\text{ext}}(t) = (H(t), 0, 0)$ of maximum intensity $H_{\max} > 0$ for 150 ps; see Figure 2.9. Then, we turn off the applied external field, i.e., $\boldsymbol{H}_{\text{ext}} \equiv (0, 0, 0)$, and let the system relax to equilibrium. In order to capture all possible excitation modes, during the application of the field and the subsequent relaxation process, we set the value of the Gilbert damping constant to $\alpha = 0.002$, which is considerably smaller than the experimental value of $\alpha = 0.28$ measured for FeGe; see [BAB$^+$17]. To probe the limit of the stability of the skyrmion, we test different values for the maximum intensity of the field $H_{\max}$, namely $\mu_0 H_{\max} = 1, 2, 5, 10, 20, 50, 100, 200$ mT. For the spatial discretization,
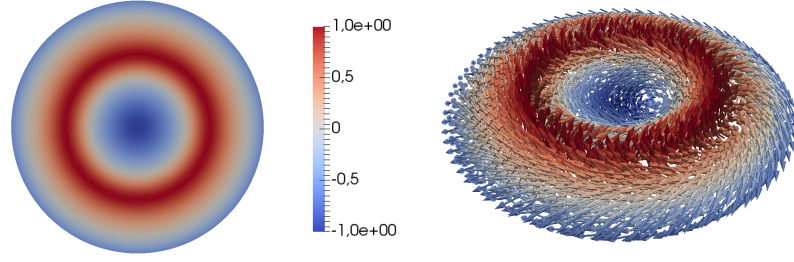
(a) Relaxed skyrmion state that we use as initial condition obtained by relaxing a uniform out-of-plane state for 1 ns.
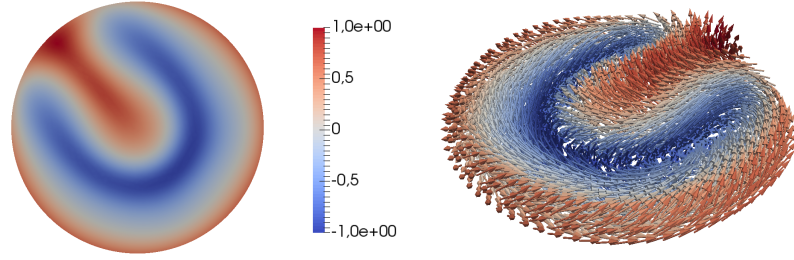
(b) Metastable horseshoe state obtained by applying a field pulse of maximum intensity $\mu_0 H_{\max} = 200$ mT to the skyrmion of (a) and relaxing the system for 10 ns.

Figure 2.8: Experiment of Section 2.4.3. Relaxed magnetization states: 2D view (left) and 3D view (right). The pictures refer to the states computed with TPS2.

we consider a regular partition of the nanodisk consisting of 36 501 tetrahedra (mesh size of 3 nm). For the time discretization, we consider a uniform time-step size of 0.1 ps.

In Figure 2.10, we plot the first 10 ns of the time evolution of the second component of the spatially averaged magnetization of the sample, i.e., $\langle m_2(t) \rangle = |\Omega|^{-1} \int_\Omega m_2(\boldsymbol{x}, t) \, \mathrm{d}\boldsymbol{x}$. We see that, for the values $\mu_0 H_{\max} = 1$, 2, 5, 10, 20, 50 mT, the induced dynamics is a periodic damped precession of the skyrmion around the center of the sample, which comes back to the initial stable configuration by the relaxation process. As expected, both the deflection from the stable symmetric initial state and the amplitude of the oscillations increase for larger values of $H_{\max}$. For the value $\mu_0 H_{\max} = 100$ mT, the skyrmion is critically deformed by the applied field pulse, but the initial stable configuration is recovered by the relaxation process. Note that a different oscillating mode comes into play in this case. For $\mu_0 H_{\max} = 200$ mT, the skyrmion is destroyed. After approximately 3.5 ns of chaotic dynamics, the magnetization configuration turns into a horseshoe state which then starts to rotate around the center of the sample; see Figure 2.8b.

As observed for the experiment of Section 2.4.2, also in this case the results obtained with TPS1 and TPS2 are in full quantitative agreement with each other. The use of PF-TPS1 preserves the qualitative outcome of the experiment, but the computed quantities, e.g., the amplitudes of the oscillations depicted in Figure 2.10, are slightly perturbed.

The presented experiment is a preliminary study to investigate the stability and the dy-
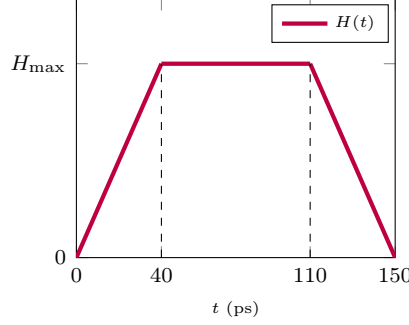
Figure 2.9: Experiment of Section 2.4.3. Structure of the applied field pulse: The field intensity increases linearly in time for 40 ps to reach the maximum value $H_{\max}$. It is then constant and equal to $H_{\max}$ for 70 ps. Finally, it decreases linearly for 40 ps to reach the value 0.

namics of a skyrmion in the presence of a field pulse. This is done to explore the possibility to use time-resolved scanning Kerr microscopy [KSH$^+$17] which is based on the interplay between laser and field pulses to directly map the dynamics of magnetic skyrmions [KSH$^+$16].

2.5 Convergence analysis

In this section, we show that all proposed algorithms are well-posed and we present the proof of Theorem 2.3.5. To establish the convergence result, we use the standard energy method for proving existence of solutions of linear second-order parabolic problems; see, e.g., [Eva10, Section 7.1.2]. The main difference is that, following [Alo08, AHP$^+$14, AKST14], the construction of approximate solutions is not obtained by applying the Galerkin method based on a basis of appropriately normalized eigenfunctions of the Laplace operator, but rather by using the finite element solutions delivered by the numerical schemes.

2.5.1 Preliminaries

We introduce some further notation and collect some auxiliary results. We consider the nodal interpolant $\mathcal{I}_h : C^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)$, which is defined by $\mathcal{I}_h[v](\boldsymbol{z}) = v(\boldsymbol{z})$ for all $\boldsymbol{z} \in \mathcal{N}_h$ and $v \in C^0(\overline{\Omega})$. It is well known that, for κ-shape-regular meshes and any integer $0 \leq m \leq 2$, the nodal interpolant satisfies the approximation property

$$\|D^m(v - \mathcal{I}_h[v])\|_{L^2(\Omega)} \leq Ch^{2-m}\|D^2v\|_{L^2(\Omega)} \quad \text{for all } v \in H^2(\Omega), \tag{2.29}$$

where the constant $C > 0$ depends only on κ. We denote the vector-valued realization of the nodal interpolant by $\boldsymbol{\mathcal{I}}_h : \boldsymbol{C}^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)^3$. The following classical inverse estimate requires the quasi-uniformity of the underlying family of meshes: For any $1 \leq p \leq \infty$, it holds that

$$\|\boldsymbol{\nabla}\boldsymbol{\phi}_h\|_{\boldsymbol{L}^p(\Omega)} \leq C_{\text{inv}}h^{-1}\|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^p(\Omega)} \quad \text{for all } \boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3, \tag{2.30}$$
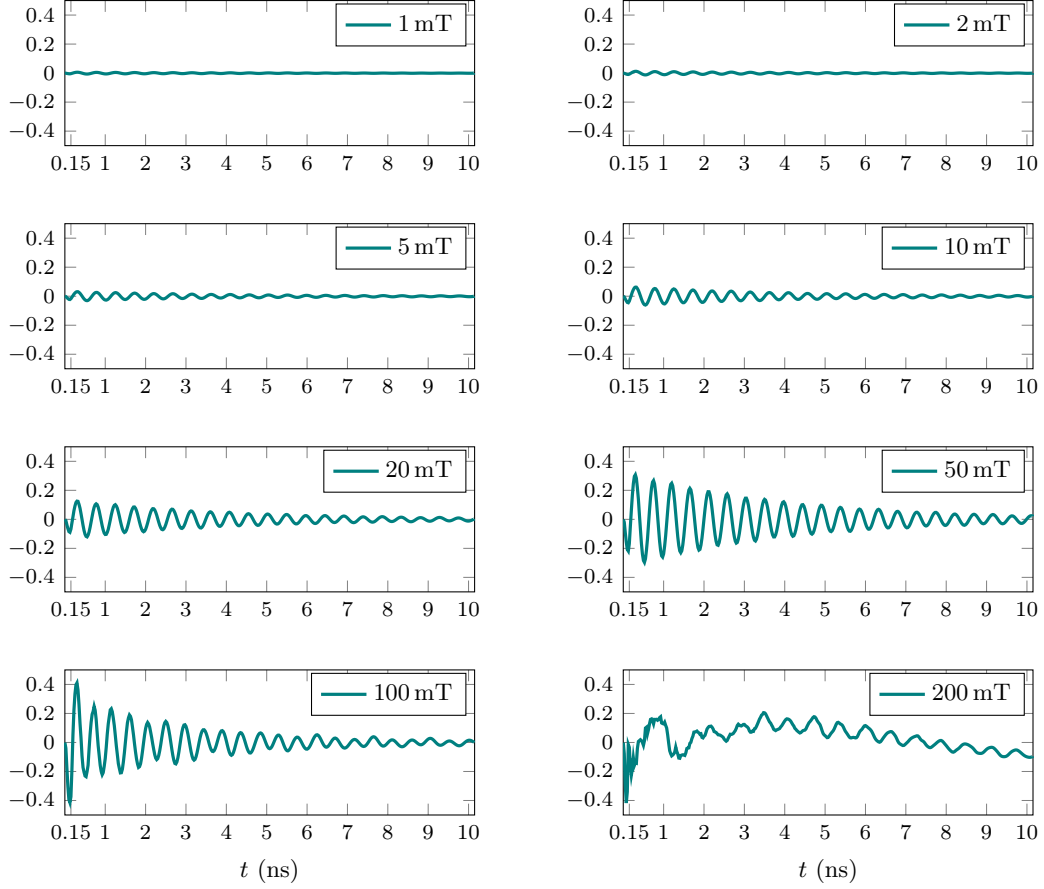
Figure 2.10: Experiment of Section 2.4.3. Time evolution of $\langle m_2 \rangle$ for applied field pulses with different intensities ranging from $\mu_0 H_{\max} = 1\,\mathrm{mT}$ to $\mu_0 H_{\max} = 200\,\mathrm{mT}$. The plots refer to the results computed with TPS2.

where $C_{\mathrm{inv}} > 0$ depends only on κ and p. Standard scaling arguments show that, for any $1 \leq p < \infty$, we have the discrete norm equivalence

$$C_{\mathrm{norm}}^{-1} \|\phi_h\|_{\boldsymbol{L}^p(\Omega)} \leq \left(h^3 \sum_{\boldsymbol{z} \in \mathcal{N}_h} |\phi_h(\boldsymbol{z})|^p \right)^{1/p} \leq C_{\mathrm{norm}} \|\phi_h\|_{\boldsymbol{L}^p(\Omega)} \quad \text{for all } \phi_h \in \mathcal{S}^1(\mathcal{T}_h)^3, \tag{2.31}$$

where $C_{\mathrm{norm}} > 0$ depends only on κ and p.

To ensure that the approximate magnetization belongs to $\boldsymbol{\mathcal{M}}_h$, TPS1 and TPS2 employ the nodal projection. We refer to [Bar05, Lemma 3.2] for the proof that the nodal projection $\phi_h \mapsto \mathcal{I}_h[\phi_h/|\phi_h|]$ does not increase the exchange energy of a discrete function if the underlying mesh fulfills a weak acuteness condition. Specifically, the angle condition (2.15) ensures that any $\phi_h \in \mathcal{S}^1(\mathcal{T}_h)^3$ with $|\phi_h(\boldsymbol{z})| \geq 1$ for all $\boldsymbol{z} \in \mathcal{N}_h$ satisfies that

$$\|\boldsymbol{\nabla} \mathcal{I}_h[\phi_h/|\phi_h|]\|_{\boldsymbol{L}^2(\Omega)} \leq \|\boldsymbol{\nabla} \phi_h\|_{\boldsymbol{L}^2(\Omega)}. \tag{2.32}$$

Owing to the application of the nodal projection, for all $0 \leq i \leq N-1$, the iterates of TPS1 and TPS2 satisfy $\|\boldsymbol{m}_h^{i+1}\|_{\boldsymbol{L}^\infty(\Omega)} = 1$ and the geometric estimates

$$|\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) - \boldsymbol{m}_h^i(\boldsymbol{z})| \leq k|\boldsymbol{v}_h^i(\boldsymbol{z})| \text{ and } |\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) - \boldsymbol{m}_h^i(\boldsymbol{z}) - k\boldsymbol{v}_h^i(\boldsymbol{z})| \leq \frac{1}{2}k^2|\boldsymbol{v}_h^i(\boldsymbol{z})|^2$$

for any $\boldsymbol{z} \in \mathcal{N}_h$; see [AJ06, BKP08]. With (2.31), these nodewise inequalities are turned into

$$\|\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i\|_{\boldsymbol{L}^p(\Omega)} \leq C_{\text{geo}}k\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^p(\Omega)}, \tag{2.33a}$$

$$\|\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i\|_{\boldsymbol{L}^p(\Omega)} \leq C_{\text{geo}}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^{2p}(\Omega)}^2, \tag{2.33b}$$

where $C_{\text{geo}} > 0$ depends only on κ and p. In the case of PF-TPS1, where the nodal projection is omitted, the geometric estimates (2.33) become trivial, but the equality $\|\boldsymbol{m}_h^{i+1}\|_{\boldsymbol{L}^\infty(\Omega)} = 1$ does not hold anymore. However, for any $1 \leq j \leq N$, the linear time-stepping yields the recursive relation

$$|\boldsymbol{m}_h^j(\boldsymbol{z})|^2 = |\boldsymbol{m}_h^0(\boldsymbol{z})|^2 + k^2 \sum_{i=0}^{j-1} |\boldsymbol{v}_h^i(\boldsymbol{z})|^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h. \tag{2.34}$$

Together with (2.31), this leads to the estimate

$$C_{\text{norm}}^{-4}\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + k^2 \sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2. \tag{2.35}$$

2.5.2 Well-posedness

In the following proposition, we prove that the three proposed algorithms are all well-posed.

Proposition 2.5.1. *Let $0 \leq i \leq N-1$. There exists a unique solution $\boldsymbol{v}_h^i \in \mathcal{K}_h(\boldsymbol{m}_h^i)$ of (2.19) and (2.20). There exists a threshold time-step size $k_0 > 0$, which depends only on α, ℓ_{ex}, and ℓ_{dm}, such that, if $k \leq k_0$, there exists a unique solution $\boldsymbol{v}_h^i \in \mathcal{K}_h(\boldsymbol{m}_h^i)$ of (2.26). The time-steppings of all algorithms are well-defined.*

Proof. For any $0 \leq i \leq N-1$, let $a_{hk}^i(\cdot, \cdot)$ be the bilinear form appearing on the left-hand side of (2.19) and (2.20). For any $\boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that

$$a_{hk}^i(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) = \alpha\|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\text{ex}}^2\theta k\|\boldsymbol{\nabla}\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)}^2.$$

Hence, the bilinear form is elliptic, even on the full space $\mathcal{S}^1(\mathcal{T}_h)^3$. Existence and uniqueness of the solution $\boldsymbol{v}_h^i \in \mathcal{K}_h(\boldsymbol{m}_h^i)$ of (2.19) and (2.20) thus follow from the Lax–Milgram theorem.

Similarly, for any $0 \leq i \leq N-1$, let $b_{hk}^i(\cdot, \cdot)$ denote the bilinear form on the left-hand side of (2.26). For any $\varepsilon > 0$ and $\boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, using (2.10) and (2.24), we deduce that

$$b_{hk}^i(\boldsymbol{\phi}_h, \boldsymbol{\phi}_h) = \langle W_{M(k)}(\lambda_h^i)\boldsymbol{\phi}_h, \boldsymbol{\phi}_h \rangle + \frac{\ell_{\text{ex}}^2}{2}k[1 + \rho(k)]\|\boldsymbol{\nabla}\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\text{dm}}}{2}k\langle\boldsymbol{\phi}_h, \nabla \times \boldsymbol{\phi}_h\rangle$$

$$\geq \left(\frac{2\alpha^2}{2\alpha + M(k)k} - \frac{\ell_{\text{dm}}}{4\varepsilon}k\right)\|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{1}{2}\left(\ell_{\text{ex}}^2 - \frac{\varepsilon\ell_{\text{dm}}}{2}\right)k\|\boldsymbol{\nabla}\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)}^2.$$

We choose $\varepsilon = \ell_{\mathrm{ex}}^2/\ell_{\mathrm{dm}}$. With the properties (2.25) of the cut-off function $M(k)$, it follows that both coefficients in front of the norms are positive if the time-step size k is sufficiently small. Then, the bilinear form is elliptic and (2.26) admits a unique solution $\boldsymbol{v}_h^i \in \boldsymbol{\mathcal{K}}_h(\boldsymbol{m}_h^i)$.

The linear time-stepping in step (ii) of PF-TPS1 is clearly well-defined. In the case of TPS1 and TPS2, which include the nodal projection, since $\boldsymbol{m}_h^i \in \boldsymbol{\mathcal{M}}_h$ and $\boldsymbol{v}_h \in \boldsymbol{\mathcal{K}}_h(\boldsymbol{m}_h^i)$, it holds that

$$|\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})|^2 = |\boldsymbol{m}_h^i(\boldsymbol{z})|^2 + k^2|\boldsymbol{v}_h^i(\boldsymbol{z})|^2 = 1 + k^2|\boldsymbol{v}_h^i(\boldsymbol{z})|^2 \geq 1 \quad \text{for any } \boldsymbol{z} \in \mathcal{N}_h.$$

The time-steppings of TPS1 and TPS2 are therefore also well defined. $\qquad\square$

2.5.3 Discrete energy law and stability

In this section, we establish the discrete energy laws and study the stability for the discrete iterates delivered by the algorithms. We first observe that, given $C_0 > 0$, assumption (2.28) provides some $h_0 > 0$ such that $\|\boldsymbol{m}_h^0\|_{\boldsymbol{H}^1(\Omega)} \leq C_0$ for all $h \leq h_0$.

In the following proposition, we prove the result for TPS1.

Proposition 2.5.2 (Discrete energy law and stability of TPS1)**.** *Let $1 \leq j \leq N$. There exists a constant $C > 0$, which depends only on κ and ℓ_{dm}, such that the iterates of TPS1 satisfy the discrete energy law*

$$\mathcal{E}(\boldsymbol{m}_h^j) + \left(\alpha - Ch^{-1}k\right) k \sum_{i=0}^{j-1} \|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2 \sum_{i=0}^{j-1} \|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \mathcal{E}(\boldsymbol{m}_h^0). \quad (2.36)$$

Moreover, there exists a constant $C' > 0$, which depends only on α, κ, and ℓ_{dm}, such that, if $h \leq h_0$ and $k \leq C'h$, the iterates of TPS1 satisfy the stability estimate

$$\|\boldsymbol{m}_h^j\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{i=0}^{j-1} \|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + (\theta - 1/2)k^2 \sum_{i=0}^{j-1} \|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \leq C'', \quad (2.37)$$

where the constant $C'' > 0$ depends only on α, C_0, κ, ℓ_{ex}, ℓ_{dm}, and $|\Omega|$.

Proof. For any $0 \leq i \leq j-1$, we test (2.19) with $\boldsymbol{\phi}_h = \boldsymbol{v}_h^i \in \boldsymbol{\mathcal{K}}_h(\boldsymbol{m}_h^i)$ to obtain the identity

$$\alpha\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2\theta k\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2$$
$$= -\ell_{\mathrm{ex}}^2\langle\boldsymbol{\nabla}\boldsymbol{m}_h^i, \boldsymbol{\nabla}\boldsymbol{v}_h^i\rangle - \frac{\ell_{\mathrm{dm}}}{2}\langle\nabla\times\boldsymbol{m}_h^i, \boldsymbol{v}_h^i\rangle - \frac{\ell_{\mathrm{dm}}}{2}\langle\nabla\times\boldsymbol{v}_h^i, \boldsymbol{m}_h^i\rangle.$$

Since the angle condition (2.15) is satisfied, we obtain that

$$\frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^{i+1}\|_{\boldsymbol{L}^2(\Omega)}^2 \overset{(2.32)}{\leq} \frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^i + k\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2$$
$$= \frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2 k\langle\boldsymbol{\nabla}\boldsymbol{m}_h^i, \boldsymbol{\nabla}\boldsymbol{v}_h^i\rangle + \frac{\ell_{\mathrm{ex}}^2}{2}k^2\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.$$
$$(2.38a)$$

Hence, it follows that

$$\frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^{i+1}\|_{\boldsymbol{L}^2(\Omega)}^2 - \frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \alpha k\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2$$
$$\leq -\frac{\ell_{\mathrm{dm}}}{2}k\langle\nabla\times\boldsymbol{m}_h^i,\boldsymbol{v}_h^i\rangle - \frac{\ell_{\mathrm{dm}}}{2}k\langle\nabla\times\boldsymbol{v}_h^i,\boldsymbol{m}_h^i\rangle.$$
(2.38b)

We obtain the energy inequality

$$\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i) + \alpha k\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2$$
$$\leq -\frac{\ell_{\mathrm{dm}}}{2}k\langle\nabla\times\boldsymbol{m}_h^i,\boldsymbol{v}_h^i\rangle - \frac{\ell_{\mathrm{dm}}}{2}k\langle\nabla\times\boldsymbol{v}_h^i,\boldsymbol{m}_h^i\rangle$$
$$+\frac{\ell_{\mathrm{dm}}}{2}\langle\nabla\times\boldsymbol{m}_h^{i+1},\boldsymbol{m}_h^{i+1}\rangle - \frac{\ell_{\mathrm{dm}}}{2}\langle\nabla\times\boldsymbol{m}_h^i,\boldsymbol{m}_h^i\rangle.$$
(2.38c)

With some simple algebraic manipulations, we rewrite the last four terms of the right-hand side (those which involve the curl operator) as

$$\langle\nabla\times\boldsymbol{m}_h^{i+1},\boldsymbol{m}_h^{i+1}\rangle - \langle\nabla\times\boldsymbol{m}_h^i,\boldsymbol{m}_h^i\rangle - k\langle\nabla\times\boldsymbol{m}_h^i,\boldsymbol{v}_h^i\rangle - k\langle\nabla\times\boldsymbol{v}_h^i,\boldsymbol{m}_h^i\rangle$$
$$= \langle\nabla\times(\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i),\boldsymbol{m}_h^{i+1}\rangle + \langle\nabla\times\boldsymbol{m}_h^i,\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i\rangle$$
$$+ k\langle\nabla\times\boldsymbol{v}_h^i,\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i\rangle.$$
(2.39)

Using the inverse estimate (2.30) and the geometric estimates (2.33), we infer that

$$|\langle\nabla\times(\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i),\boldsymbol{m}_h^{i+1}\rangle| \leq \|\nabla\times(\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i)\|_{\boldsymbol{L}^1(\Omega)}\|\boldsymbol{m}_h^{i+1}\|_{\boldsymbol{L}^\infty(\Omega)}$$
$$\leq \sqrt{2}\,C_{\mathrm{inv}}h^{-1}\|\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i\|_{\boldsymbol{L}^1(\Omega)}$$
$$\leq \sqrt{2}\,C_{\mathrm{inv}}C_{\mathrm{geo}}h^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.$$

Similarly, it holds that

$$|\langle\nabla\times\boldsymbol{m}_h^i,\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i\rangle| \leq \|\nabla\times\boldsymbol{m}_h^i\|_{\boldsymbol{L}^\infty(\Omega)}\|\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i\|_{\boldsymbol{L}^1(\Omega)}$$
$$\leq \sqrt{2}\,C_{\mathrm{inv}}C_{\mathrm{geo}}h^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2$$

and

$$k|\langle\nabla\times\boldsymbol{v}_h^i,\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i\rangle| \leq k\|\nabla\times\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i\|_{\boldsymbol{L}^2(\Omega)}$$
$$\leq \sqrt{2}\,C_{\mathrm{inv}}C_{\mathrm{geo}}h^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.$$

With $C = 3\,C_{\mathrm{inv}}C_{\mathrm{geo}}\ell_{\mathrm{dm}}/\sqrt{2}$, we thus obtain that

$$\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i) + \alpha k\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \leq Ch^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.$$

Summation over $0 \leq i \leq j - 1$ leads to (2.36).

To show (2.37), we first note that $\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^\infty(\Omega)} = 1$ yields that

$$\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 \leq |\Omega|.$$
(2.40)

We multiply (2.40) by $\ell_{\mathrm{dm}}^2/\ell_{\mathrm{ex}}^2$ and add the resulting equation to (2.36). Using the characterization (2.11) of the energy, we obtain that

$$
\frac{\ell_{\mathrm{ex}}^2}{4}\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\mathrm{dm}}^2}{2\ell_{\mathrm{ex}}^2}\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 + \left(\alpha - Ch^{-1}k\right)k\sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2
$$
$$
+ \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2\sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \le \frac{\ell_{\mathrm{ex}}^2 + \ell_{\mathrm{dm}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{1}{4}\|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\mathrm{dm}}^2}{\ell_{\mathrm{ex}}^2}|\Omega|.
$$

Let $C' = \alpha/(2C)$ and $k \le C'h$. Since $\theta \ge 1/2$, all terms on the left-hand side are nonnegative. We obtain (2.37), where the constant $C'' > 0$ (which we do not compute explicitly) depends only on α, C_0, κ, ℓ_{ex}, ℓ_{dm}, and $|\Omega|$. $\qquad\square$

In the following proposition, we prove the corresponding result for PF-TPS1.

Proposition 2.5.3 (Discrete energy law and stability of PF-TPS1). *Let $1 \le j \le N$ and $1/2 < \theta \le 1$. The iterates of PF-TPS1 satisfy the discrete energy law*

$$
\mathcal{E}(\boldsymbol{m}_h^j) + \alpha k\sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2\sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2
$$
$$
= \mathcal{E}(\boldsymbol{m}_h^0) + \frac{\ell_{\mathrm{dm}}}{2}k^2\sum_{i=0}^{j-1}\langle\nabla\times\boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle.
$$

<div align="right">(2.41)</div>

Moreover, there exists a threshold time-step size $k_0 > 0$, which depends only on α, κ, ℓ_{ex}, ℓ_{dm}, and θ, such that, if $h \le h_0$ and $k \le k_0$, the iterates of PF-TPS1 satisfy the stability estimate

$$
\|\boldsymbol{m}_h^j\|_{\boldsymbol{H}^1(\Omega)}^2 + k\sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + k^2\sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \le C,
\tag{2.42}
$$

where the constant $C > 0$ depends only on α, C_0, κ, ℓ_{ex}, ℓ_{dm}, and θ.

Proof. Let $0 \le i \le j - 1$. We follow the argument of the proof of Proposition 2.5.2: Due to the linear time-stepping of PF-TPS1, all the computations in (2.38) hold with equality sign and without resorting to the angle condition (2.15). Moreover, all but the last term on the right-hand side of (2.39) vanish. As a result, we obtain the energy identity

$$
\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i) + \alpha k\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 = \frac{\ell_{\mathrm{dm}}}{2}k^2\langle\nabla\times\boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle.
$$

Summing this identity over $0 \le i \le j - 1$, we prove (2.41). To estimate the right-hand side, we apply the weighted Young inequality (2.10), which, for any $\varepsilon > 0$, yields that

$$
\langle\nabla\times\boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle \le \sqrt{2}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)} \le \varepsilon\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{1}{2\varepsilon}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.
$$

Choosing $\varepsilon = \ell_{\mathrm{ex}}^2(\theta - 1/2)/\ell_{\mathrm{dm}}$, the characterization (2.11) of the energy shows that

$$
\begin{aligned}
&\frac{\ell_{\mathrm{ex}}^2}{4}\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 - \frac{\ell_{\mathrm{dm}}^2}{2\ell_{\mathrm{ex}}^2}\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 + \left(\alpha - \frac{\ell_{\mathrm{dm}}^2}{2\ell_{\mathrm{ex}}^2(2\theta - 1)}k\right) k \sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \\
&\quad + \frac{\ell_{\mathrm{ex}}^2}{2}(\theta - 1/2)k^2 \sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \le \frac{\ell_{\mathrm{ex}}^2 + \ell_{\mathrm{dm}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{1}{4}\|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2.
\end{aligned}
\tag{2.43}
$$

We multiply (2.35) by $C_{\mathrm{norm}}^4\ell_{\mathrm{dm}}^2/\ell_{\mathrm{ex}}^2$ and add the resulting equation to (2.43) to obtain that

$$
\begin{aligned}
&\frac{\ell_{\mathrm{ex}}^2}{4}\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\mathrm{dm}}^2}{2\ell_{\mathrm{ex}}^2}\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 + \left(\alpha - \frac{\ell_{\mathrm{dm}}^2[1 + 2C_{\mathrm{norm}}^4(2\theta - 1)]}{2\ell_{\mathrm{ex}}^2(2\theta - 1)}k\right) k \sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \\
&\quad + \frac{\ell_{\mathrm{ex}}^2}{2}(\theta - 1/2)k^2 \sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \\
&\qquad \le \frac{\ell_{\mathrm{ex}}^2 + \ell_{\mathrm{dm}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\mathrm{ex}}^2 + 4C_{\mathrm{norm}}^4\ell_{\mathrm{dm}}^2}{4\ell_{\mathrm{ex}}^2}\|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2.
\end{aligned}
$$

If $k \le k_0 = \alpha\,\ell_{\mathrm{ex}}^2(2\theta - 1)/\{\ell_{\mathrm{dm}}^2[1 + 2C_{\mathrm{norm}}^4(2\theta - 1)]\}$, then all terms on the left-hand side are nonnegative. This leads to (2.42), where the (explicitly computable) constant $C > 0$ depends only on α, C_0, κ, ℓ_{ex}, ℓ_{dm}, and θ. $\qquad\square$

Finally, in the following proposition, we prove the stability result for TPS2.

Proposition 2.5.4 (Discrete energy law and stability of TPS2). *Let $1 \le j \le N$. Suppose that the time-step size k is sufficiently small, so that (2.26) is well-posed by Proposition 2.5.1. Then, there exists a constant $C > 0$, which depends only on κ and ℓ_{dm}, such that the iterates of TPS2 satisfy the discrete energy law*

$$
\begin{aligned}
&\mathcal{E}(\boldsymbol{m}_h^j) + k \sum_{i=0}^{j-1}\langle W_{M(k)}(\lambda_h^i)\boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle + \frac{\ell_{\mathrm{ex}}^2}{2}\rho(k)k^2 \sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \\
&\qquad \le \mathcal{E}(\boldsymbol{m}_h^0) + Ch^{-1}k^2 \sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.
\end{aligned}
\tag{2.44}
$$

Moreover, there exists a threshold time-step size $k_0 > 0$, which depends only on α, ℓ_{ex}, and ℓ_{dm}, and a constant $C' > 0$, which depends only on α, κ, and ℓ_{dm}, such that, if $h \le h_0$, $k \le k_0$, and $k \le C'h$, the iterates of TPS2 satisfy the stability estimate

$$
\|\boldsymbol{m}_h^j\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 + \rho(k)k^2 \sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \le C'',
\tag{2.45}
$$

where the constant $C'' > 0$ depends only on α, C_0, κ, ℓ_{ex}, ℓ_{dm}, and $|\Omega|$.

Proof. Let $0 \leq i \leq j - 1$. We follow step by step the argument of the proof of Proposition 2.5.2 to obtain the inequality

$$\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i) + k\langle W_{M(k)}(\lambda_h^i)\boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle + \frac{\ell_{\text{ex}}^2}{2}\rho(k)k^2\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2$$

$$\leq -\frac{\ell_{\text{dm}}}{2}k^2\langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle - \frac{\ell_{\text{dm}}}{2}k\langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{v}_h^i\rangle - \frac{\ell_{\text{dm}}}{2}k\langle \boldsymbol{m}_h^i, \nabla \times \boldsymbol{v}_h^i\rangle$$

$$+ \frac{\ell_{\text{dm}}}{2}\langle \nabla \times \boldsymbol{m}_h^{i+1}, \boldsymbol{m}_h^{i+1}\rangle - \frac{\ell_{\text{dm}}}{2}\langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{m}_h^i\rangle.$$

We reformulate the terms of the right-hand side which involve the curl operator, i.e.,

$$\langle \nabla \times \boldsymbol{m}_h^{i+1}, \boldsymbol{m}_h^{i+1}\rangle - \langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{m}_h^i + k\boldsymbol{v}_h^i\rangle - k\langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{m}_h^i\rangle - k^2\langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle$$

$$= \langle \nabla \times (\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i), \boldsymbol{m}_h^{i+1}\rangle + \langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i\rangle$$

$$+ k\langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i\rangle - k^2\langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle$$

and proceed with their direct estimation: Using (2.30) and (2.33), we obtain that

$$|\langle \nabla \times (\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i), \boldsymbol{m}_h^{i+1}\rangle| \leq \sqrt{2}\,C_{\text{inv}}C_{\text{geo}}h^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2,$$

$$|\langle \nabla \times \boldsymbol{m}_h^i, \boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i - k\boldsymbol{v}_h^i\rangle| \leq \sqrt{2}\,C_{\text{inv}}C_{\text{geo}}h^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2,$$

$$k|\langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i\rangle| \leq \sqrt{2}\,C_{\text{inv}}C_{\text{geo}}h^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2,$$

$$k^2|\langle \nabla \times \boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle| \leq \sqrt{2}\,C_{\text{inv}}h^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.$$

It follows that

$$\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i) + k\langle W_{M(k)}(\lambda_h^i)\boldsymbol{v}_h^i, \boldsymbol{v}_h^i\rangle + \frac{\ell_{\text{ex}}^2}{2}\rho(k)k^2\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \leq Ch^{-1}k^2\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2,$$

with $C = C_{\text{inv}}(3C_{\text{geo}} + 1)\ell_{\text{dm}}/\sqrt{2}$. Summation over $0 \leq i \leq j - 1$ leads to (2.44).

If k is sufficiently small, it holds that $W_{M(k)}(\cdot) \geq \alpha/2$; see (2.24)–(2.25). With the characterization (2.11) of the energy and the inequality $\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 \leq |\Omega|$, we obtain that

$$\frac{\ell_{\text{ex}}^2}{4}\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\text{dm}}^2}{2\ell_{\text{ex}}^2}\|\boldsymbol{m}_h^j\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\alpha - 2Ch^{-1}k}{2}k\sum_{i=0}^{j-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2$$

$$+ \frac{\ell_{\text{ex}}^2}{2}\rho(k)k^2\sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \frac{\ell_{\text{ex}}^2 + \ell_{\text{dm}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{1}{4}\|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\ell_{\text{dm}}^2}{\ell_{\text{ex}}^2}|\Omega|.$$

Let $C' = \alpha/(4C)$. If $k \leq C'h$, then all terms on the left-hand side are nonnegative. Hence, we obtain (2.45), where the constant $C'' > 0$ (which we do not compute explicitly) depends only on α, C_0, κ, ℓ_{ex}, ℓ_{dm}, and $|\Omega|$. $\qquad\square$

2.5.4 Extraction of weakly convergent subsequences

Exploiting the established stability estimates of the three algorithms, we are now able to prove that the time reconstructions defined by (2.27) are uniformly bounded.

Proposition 2.5.5. *Suppose that the assumptions of Theorem 2.3.5 are satisfied. For any algorithm, if h and k are sufficiently small, the sequences $\{\boldsymbol{m}_{hk}\}$, $\{\boldsymbol{m}_{hk}^{\pm}\}$, and $\{\boldsymbol{v}_{hk}^{-}\}$ are uniformly bounded in the sense that*

$$\|\boldsymbol{m}_{hk}\|_{L^\infty(0,T;\boldsymbol{H}^1(\Omega))} + \|\boldsymbol{m}_{hk}^{\pm}\|_{L^\infty(0,T;\boldsymbol{H}^1(\Omega))} + \|\partial_t\boldsymbol{m}_{hk}\|_{\boldsymbol{L}^2(\Omega_T)} + \|\boldsymbol{v}_{hk}^{-}\|_{\boldsymbol{L}^2(\Omega_T)} \le C. \quad (2.46)$$

The constant $C > 0$ is independent of h and k. Moreover, it holds that

$$\lim_{h,k\to 0} k\|\boldsymbol{\nabla}\boldsymbol{v}_{hk}^{-}\|_{\boldsymbol{L}^2(\Omega_T)} = 0. \quad (2.47)$$

Proof. The estimate (2.46) follows directly from (2.37), (2.42), and (2.45). For TPS1 and TPS2, also the geometric estimate (2.33a) is used to conclude that $\|\partial_t\boldsymbol{m}_{hk}\|_{\boldsymbol{L}^2(\Omega_T)} \le C$.

The convergence (2.47) for PF-TPS1 follows from Proposition 2.5.3. Indeed, it holds that

$$k^2\|\boldsymbol{\nabla}\boldsymbol{v}_{hk}^{-}\|_{\boldsymbol{L}^2(\Omega_T)}^2 = k^3\sum_{i=0}^{N-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \overset{(2.42)}{\le} Ck.$$

For TPS1 (resp. TPS2), we first resort to an inverse estimate to obtain that

$$k^2\|\boldsymbol{\nabla}\boldsymbol{v}_{hk}^{-}\|_{\boldsymbol{L}^2(\Omega_T)}^2 = k^3\sum_{i=0}^{N-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \overset{(2.30)}{\le} C_{\mathrm{inv}}h^{-2}k^3\sum_{i=0}^{N-1}\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.$$

The result then follows from Proposition 2.5.2 (resp. Proposition 2.5.4) and the fact that $k/h \to 0$ as $h, k \to 0$ by assumption. $\qquad\square$

With this result, we can now extract weakly convergent subsequences.

Proposition 2.5.6. *Suppose that the assumptions of Theorem 2.3.5 are satisfied. Then, for any algorithm, there exists $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T) \cap L^\infty(0,T;\boldsymbol{H}^1(\Omega))$, which satisfies $|\boldsymbol{m}| = 1$ a.e. in Ω_T, such that the sequences of time reconstructions $\{\boldsymbol{m}_{hk}\}$, $\{\boldsymbol{m}_{hk}^{\pm}\}$, and $\{\boldsymbol{v}_{hk}^{-}\}$ admit subsequences (not relabeled) for which it holds that*

$$\boldsymbol{m}_{hk} \rightharpoonup \boldsymbol{m} \quad \text{in } \boldsymbol{H}^1(\Omega_T), \quad (2.48a)$$

$$\boldsymbol{m}_{hk} \to \boldsymbol{m} \quad \text{in } \boldsymbol{H}^s(\Omega_T) \text{ for all } 0 < s < 1, \quad (2.48b)$$

$$\boldsymbol{m}_{hk} \to \boldsymbol{m} \quad \text{in } L^2(0,T;\boldsymbol{H}^s(\Omega)) \text{ for all } 0 < s < 1, \quad (2.48c)$$

$$\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^{\pm} \to \boldsymbol{m} \quad \text{in } \boldsymbol{L}^2(\Omega_T), \quad (2.48d)$$

$$\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^{\pm} \to \boldsymbol{m} \quad \text{pointwise a.e. in } \Omega_T, \quad (2.48e)$$

$$\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^{\pm} \overset{\star}{\rightharpoonup} \boldsymbol{m} \quad \text{in } L^\infty(0,T;\boldsymbol{H}^1(\Omega)), \quad (2.48f)$$

$$\boldsymbol{v}_{hk}^{-} \rightharpoonup \partial_t\boldsymbol{m} \quad \text{in } \boldsymbol{L}^2(\Omega_T) \quad (2.48g)$$

as $h, k \to 0$.

Proof. For the sake of clarity, we divide the proof into three steps.

- **Step 1:** Proof of the convergence results (2.48a)–(2.48f).

The uniform boundedness (2.46) established by Proposition 2.5.5 allows us to extract weakly convergent subsequences (not relabeled) of $\{\boldsymbol{m}_{hk}\}$, $\{\boldsymbol{m}_{hk}^{\pm}\}$, with possibly different limits, and $\{\boldsymbol{v}_{hk}^{-}\}$ in $\boldsymbol{H}^1(\Omega_T)$, $L^2(0,T;\boldsymbol{H}^1(\Omega))$, and $\boldsymbol{L}^2(\Omega_T)$, respectively.

Let $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T)$ denote the weak limit of $\{\boldsymbol{m}_{hk}\}$ in $\boldsymbol{H}^1(\Omega_T)$. The continuous inclusions $\boldsymbol{H}^1(\Omega_T) \subset L^2(0,T;\boldsymbol{H}^1(\Omega)) \subset \boldsymbol{L}^2(\Omega_T)$ and the compact embedding $\boldsymbol{H}^1(\Omega_T) \Subset \boldsymbol{L}^2(\Omega_T)$ show that $\boldsymbol{m}_{hk} \rightharpoonup \boldsymbol{m}$ in $L^2(0,T;\boldsymbol{H}^1(\Omega))$ and $\boldsymbol{m}_{hk} \to \boldsymbol{m}$ in $\boldsymbol{L}^2(\Omega_T)$. In particular, upon extraction of a further subsequence, we obtain that $\boldsymbol{m}_{hk} \to \boldsymbol{m}$ pointwise a.e. in Ω_T.

Let $0 < s < 1$. From the interpolation result $[\boldsymbol{L}^2(\Omega_T), \boldsymbol{H}^1(\Omega_T)]_s = \boldsymbol{H}^s(\Omega_T)$, we obtain the compact embedding $\boldsymbol{H}^1(\Omega_T) \Subset \boldsymbol{H}^s(\Omega_T)$; see, e.g., [BL76, Theorem 6.4.5 and Theorem 3.8.1]. Since $[\boldsymbol{L}^2(\Omega_T), L^2(0,T;\boldsymbol{H}^1(\Omega))]_s = L^2(0,T;\boldsymbol{H}^s(\Omega))$, which follows, e.g., by [BL76, Theorem 5.1.2], we deduce that the inclusion $\boldsymbol{H}^s(\Omega_T) \subset L^2(0,T;\boldsymbol{H}^s(\Omega))$ is continuous. Hence, it holds that $\boldsymbol{H}^1(\Omega_T) \Subset \boldsymbol{H}^s(\Omega_T) \subset L^2(0,T;\boldsymbol{H}^s(\Omega))$, from which we conclude that $\boldsymbol{m}_{hk} \to \boldsymbol{m}$ in both $\boldsymbol{H}^s(\Omega_T)$ and $L^2(0,T;\boldsymbol{H}^s(\Omega))$. Moreover, since

$$\|\boldsymbol{m}_{hk} - \boldsymbol{m}_{hk}^{\pm}\|_{\boldsymbol{L}^2(\Omega_T)} \overset{(2.27)}{\leq} k\|\partial_t \boldsymbol{m}_{hk}\|_{\boldsymbol{L}^2(\Omega_T)} \overset{(2.46)}{\lesssim} k,$$

it follows that $\boldsymbol{m}_{hk}^{\pm} \rightharpoonup \boldsymbol{m}$ in $L^2(0,T;\boldsymbol{H}^1(\Omega))$ as well as $\boldsymbol{m}_{hk}^{\pm} \to \boldsymbol{m}$ in $\boldsymbol{L}^2(\Omega_T)$ and pointwise a.e. in Ω_T. Finally, since the sequences $\{\boldsymbol{m}_{hk}\}$ and $\{\boldsymbol{m}_{hk}^{\pm}\}$ are uniformly bounded also in $L^\infty(0,T;\boldsymbol{H}^1(\Omega))$, we can extract further weakly-star convergent subsequences, whose limits coincide with the weak limits in $L^2(0,T;\boldsymbol{H}^1(\Omega))$, i.e., it holds that $\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^{\pm} \overset{\star}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(0,T;\boldsymbol{H}^1(\Omega))$.

- **Step 2:** Proof of (2.48g).

Let $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega_T)$ such that $\boldsymbol{v}_{hk}^{-} \rightharpoonup \boldsymbol{v}$ in $\boldsymbol{L}^2(\Omega_T)$. In the case of TPS1 and TPS2, which includes the nodal projection, it holds that

$$\|\partial_t \boldsymbol{m} - \boldsymbol{v}\|_{\boldsymbol{L}^1(\Omega_T)} \leq \liminf_{h,k\to 0}\|\partial_t \boldsymbol{m}_{hk} - \boldsymbol{v}_{hk}^{-}\|_{\boldsymbol{L}^1(\Omega_T)} \overset{(2.33b)}{\leq} C_{\text{geo}}^2 k\|\boldsymbol{v}_{hk}^{-}\|_{\boldsymbol{L}^2(\Omega_T)}^2 \overset{(2.46)}{\lesssim} k,$$

which shows that $\boldsymbol{v} = \partial_t \boldsymbol{m}$ a.e. in Ω_T. For PF-TPS1, the result directly follows from the equality $\boldsymbol{m}_h^{i+1} = \boldsymbol{m}_h^i + k\boldsymbol{v}_h^i$.

- **Step 3:** $\boldsymbol{m}$ satisfies $|\boldsymbol{m}| = 1$ a.e. in Ω_T.

In the case of TPS1 and TPS2, since $\mathcal{I}_h[|\boldsymbol{m}_h^i|^2] = 1$ and $\boldsymbol{\nabla}\boldsymbol{m}_h^i$ is piecewise constant for all $0 \leq i \leq N-1$, it holds that

$$\big\||\boldsymbol{m}_{hk}^{-}|^2 - 1\big\|_{L^2(\Omega_T)} \lesssim h\|\boldsymbol{\nabla}\boldsymbol{m}_{hk}^{-}\|_{\boldsymbol{L}^2(\Omega_T)} \overset{(2.46)}{\lesssim} h,$$

which yields the convergence $|\boldsymbol{m}_{hk}^{-}|^2 \to 1$ in $L^2(\Omega_T)$. Since $\boldsymbol{m}_{hk}^{-} \to \boldsymbol{m}$ pointwise a.e. in Ω_T, we deduce that $|\boldsymbol{m}| = 1$ a.e. in Ω_T.

In the case of PF-TPS1, we start with a triangle inequality, which shows that

$$\big\||\boldsymbol{m}|^2 - 1\big\|_{L^1(\Omega_T)}$$
$$\leq \big\||\boldsymbol{m}|^2 - |\boldsymbol{m}_{hk}^{+}|^2\big\|_{L^1(\Omega_T)} + \big\||\boldsymbol{m}_{hk}^{+}|^2 - \mathcal{I}_h[|\boldsymbol{m}_{hk}^{+}|^2]\big\|_{L^1(\Omega_T)} + \big\|\mathcal{I}_h[|\boldsymbol{m}_{hk}^{+}|^2] - 1\big\|_{L^1(\Omega_T)}.$$

The first two terms on the right-hand side converge to 0. Indeed, on the one hand, it holds that

$$\big\| |\boldsymbol{m}|^2 - |\boldsymbol{m}_{hk}^+|^2 \big\|_{L^1(\Omega_T)} \leq \|\boldsymbol{m} + \boldsymbol{m}_{hk}^+\|_{\boldsymbol{L}^2(\Omega_T)} \|\boldsymbol{m} - \boldsymbol{m}_{hk}^+\|_{\boldsymbol{L}^2(\Omega_T)} \lesssim \|\boldsymbol{m} - \boldsymbol{m}_{hk}^+\|_{\boldsymbol{L}^2(\Omega_T)}$$

and $\boldsymbol{m}_{hk}^+ \to \boldsymbol{m}$ in $\boldsymbol{L}^2(\Omega_T)$. On the other hand, using the approximation properties (2.29) of the nodal interpolant and the fact that $\boldsymbol{\nabla}\boldsymbol{m}_h^{i+1}$ is piecewise constant, one shows that

$$\big\| |\boldsymbol{m}_{hk}^+|^2 - \mathcal{I}_h\big[|\boldsymbol{m}_{hk}^+|^2\big] \big\|_{L^1(\Omega_T)} \lesssim h^2 \|\boldsymbol{\nabla}\boldsymbol{m}_{hk}^+\|_{\boldsymbol{L}^2(\Omega_T)}^2 \overset{(2.46)}{\lesssim} h^2.$$

To conclude, it remains to show that

$$\big\| \mathcal{I}_h\big[|\boldsymbol{m}_{hk}^+|^2\big] - 1 \big\|_{L^1(\Omega_T)} \to 0. \tag{2.49}$$

For any $t \in (0, T)$, let $0 \leq i \leq N - 1$ such that $t \in [t_i, t_{i+1})$. It holds that

$$\begin{aligned}
\big\| \mathcal{I}_h\big[|\boldsymbol{m}_{hk}^+(t)|^2\big] - 1 \big\|_{L^1(\Omega)} &= \big\| \mathcal{I}_h\big[|\boldsymbol{m}_h^{i+1}|^2\big] - 1 \big\|_{L^1(\Omega)} \\
&\leq \big\| \mathcal{I}_h\big[|\boldsymbol{m}_h^{i+1}|^2\big] - \mathcal{I}_h\big[|\boldsymbol{m}_h^0|^2\big] \big\|_{L^1(\Omega)} + \big\| \mathcal{I}_h\big[|\boldsymbol{m}_h^0|^2\big] - |\boldsymbol{m}_h^0|^2 \big\|_{L^1(\Omega)} + \big\| |\boldsymbol{m}_h^0|^2 - 1 \big\|_{L^1(\Omega)}.
\end{aligned}$$

For the first term on the right-hand side, it holds that

$$\begin{aligned}
\big\| \mathcal{I}_h\big[|\boldsymbol{m}_h^{i+1}|^2\big] - \mathcal{I}_h\big[|\boldsymbol{m}_h^0|^2\big] \big\|_{L^1(\Omega)} &\overset{(2.31)}{\lesssim} h^3 \sum_{\boldsymbol{z} \in \mathcal{N}_h} \big| |\boldsymbol{m}_h^{i+1}(\boldsymbol{z})|^2 - |\boldsymbol{m}_h^0(\boldsymbol{z})|^2 \big| \\
&\overset{(2.34)}{=} h^3 \sum_{\boldsymbol{z} \in \mathcal{N}_h} k^2 \sum_{\ell=0}^{i} |\boldsymbol{v}_h^\ell(\boldsymbol{z})|^2 \\
&\overset{(2.31)}{\lesssim} k^2 \sum_{\ell=0}^{i} \|\boldsymbol{v}_h^\ell\|_{\boldsymbol{L}^2(\Omega)}^2 \overset{(2.42)}{\leq} Ck.
\end{aligned} \tag{2.50}$$

Using the approximation properties of $\mathcal{I}_h$, we estimate the second term by

$$\big\| \mathcal{I}_h\big[|\boldsymbol{m}_h^0|^2\big] - |\boldsymbol{m}_h^0|^2 \big\|_{L^1(\Omega)} \lesssim h^2 \|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 \overset{(2.28)}{\lesssim} h^2.$$

Finally, since $|\boldsymbol{m}^0| = 1$ a.e. in Ω by assumption, the third term satisfies that

$$\begin{aligned}
\big\| |\boldsymbol{m}_h^0|^2 - 1 \big\|_{L^1(\Omega)} &= \big\| |\boldsymbol{m}_h^0|^2 - |\boldsymbol{m}^0|^2 \big\|_{L^1(\Omega)} \leq \|\boldsymbol{m}_h^0 + \boldsymbol{m}^0\|_{\boldsymbol{L}^2(\Omega)} \|\boldsymbol{m}_h^0 - \boldsymbol{m}^0\|_{\boldsymbol{L}^2(\Omega)} \\
&\lesssim \|\boldsymbol{m}_h^0 - \boldsymbol{m}^0\|_{\boldsymbol{L}^2(\Omega)}.
\end{aligned}$$

Thanks to (2.28), this yields the convergence $|\boldsymbol{m}_h^0|^2 \to 1$ in $L^1(\Omega)$. Altogether, this proves (2.49) and thus concludes the proof. $\qquad\square$

2.5.5 Identification of the limit with a weak solution of LLG

We start with establishing an auxiliary convergence result for the time reconstructions obtained by PF-TPS1.

Lemma 2.5.7. *Suppose that the assumptions of Theorem 2.3.5 are satisfied and let $\{\boldsymbol{m}_{hk}^{\pm}\}$ be the time reconstructions generated by PF-TPS1. For all $0 < s < 1$, it holds that*

$$\boldsymbol{m}_{hk}^{\pm} \to \boldsymbol{m} \quad in \ L^2(0, T; \boldsymbol{H}^s(\Omega)) \quad as \ h, k \to 0.$$

Proof. Let $0 < s < 1$. It holds that

$$
\begin{aligned}
\|\boldsymbol{m}_{hk} &- \boldsymbol{m}_{hk}^{\pm}\|_{L^2(0,T;\boldsymbol{H}^s(\Omega))}^2 \\
&= \int_0^T \|\boldsymbol{m}_{hk}(t) - \boldsymbol{m}_{hk}^{\pm}(t)\|_{\boldsymbol{H}^s(\Omega)}^2 \mathrm{d}t \overset{(2.27)}{\leq} k^2 \int_0^T \|\partial_t \boldsymbol{m}_{hk}(t)\|_{\boldsymbol{H}^s(\Omega)}^2 \mathrm{d}t \\
&= k^2 \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|\partial_t \boldsymbol{m}_{hk}(t)\|_{\boldsymbol{H}^s(\Omega)}^2 \mathrm{d}t = k^3 \sum_{i=0}^{N-1} \|(\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i)/k\|_{\boldsymbol{H}^s(\Omega)}^2 \\
&= k^3 \sum_{i=0}^{N-1} \|\boldsymbol{v}_h^i\|_{\boldsymbol{H}^s(\Omega)}^2 \lesssim k^3 \sum_{i=0}^{N-1} \|\boldsymbol{v}_h^i\|_{\boldsymbol{H}^1(\Omega)}^2 = k^2 \|\boldsymbol{v}_{hk}^-\|_{\boldsymbol{L}^2(\Omega_T)}^2 + k^2 \|\boldsymbol{\nabla}\boldsymbol{v}_{hk}^-\|_{\boldsymbol{L}^2(\Omega_T)}^2.
\end{aligned}
$$

Since $\boldsymbol{m}_{hk} \to \boldsymbol{m}$ in $L^2(0, T; \boldsymbol{H}^s(\Omega))$ by Proposition 2.5.6, the result follows from (2.46)–(2.47). $\qquad\square$

We have collected all ingredients to finalize the proof of Theorem 2.3.5.

Proof of Theorem 2.3.5. By Proposition 2.5.6, for any algorithm, we deduce the desired convergence towards a function $\boldsymbol{m} \in L^\infty(0, T; \boldsymbol{H}^1(\Omega)) \cap H^1(0, T; \boldsymbol{L}^2(\Omega))$ satisfying $|\boldsymbol{m}| = 1$ a.e. in Ω_T. Since $\boldsymbol{m}_{hk} \rightharpoonup \boldsymbol{m}$ in $\boldsymbol{H}^1(\Omega_T)$, we also have the weak convergence of the traces, i.e., $\boldsymbol{m}_{hk}(0) \rightharpoonup \boldsymbol{m}(0)$ in $\boldsymbol{H}^{1/2}(\Omega)$. By assumption (2.28), we deduce that $\boldsymbol{m}(0) = \boldsymbol{m}^0$ in the sense of traces. It remains to show that $\boldsymbol{m}$ fulfills the variational formulation (2.13) and the energy inequality (2.14). For the sake of clarity, we consider the three algorithms separately.

- **Step 1:** Proof of the result for TPS1.

Let $\boldsymbol{\varphi} \in \boldsymbol{C}^\infty(\overline{\Omega_T})$ be an arbitrary test function. For any $0 \leq i \leq N-1$ and $t \in (t_i, t_{i+1})$, we test (2.19) with $\boldsymbol{\phi}_h = \mathcal{I}_h[\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \in \mathcal{K}_h(\boldsymbol{m}_h^i)$. Integrating in time over $t \in (t_i, t_{i+1})$, summing over $0 \leq i \leq N-1$, and using the approximation property (2.29) of the nodal interpolant, we obtain the identity

$$
\begin{aligned}
\alpha \int_0^T &\langle \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)\rangle \, \mathrm{d}t + \int_0^T \langle \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)\rangle \, \mathrm{d}t \\
&+ \ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{\nabla}[\boldsymbol{m}_{hk}^-(t) + \theta k \boldsymbol{v}_{hk}^-(t)], \boldsymbol{\nabla}[\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)]\rangle \, \mathrm{d}t + \mathcal{O}(h) \\
&= -\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times \boldsymbol{m}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)\rangle \, \mathrm{d}t \\
&\quad - \frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}_{hk}^-(t), \nabla \times [\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)]\rangle \, \mathrm{d}t.
\end{aligned}
\tag{2.51}
$$

Using the available convergence results, we would like to pass the latter to the limit as $h, k \to 0$ to obtain (2.13). For the left-hand side, it holds that

$$\alpha \int_0^T \langle \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \to -\alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t,$$

$$\int_0^T \langle \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \to \int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t,$$

$$\ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{\nabla}[\boldsymbol{m}_{hk}^-(t) + \theta k \boldsymbol{v}_{hk}^-(t)], \boldsymbol{\nabla} \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \to \ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{\nabla} \boldsymbol{m}(t), \boldsymbol{\nabla} \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t;$$

see [Alo08, BFF$^+$14] for details. For the first term on the right-hand side, since $\nabla \times \boldsymbol{m}_{hk}^- \rightharpoonup \nabla \times \boldsymbol{m}$ and $\boldsymbol{m}_{hk}^- \times \boldsymbol{\varphi} \to \boldsymbol{m} \times \boldsymbol{\varphi}$ in $\boldsymbol{L}^2(\Omega_T)$, it holds that

$$-\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times \boldsymbol{m}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \to -\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times \boldsymbol{m}(t), \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t.$$

Since $\nabla \times (\boldsymbol{m}_{hk}^- \times \boldsymbol{\varphi}) \rightharpoonup \nabla \times (\boldsymbol{m} \times \boldsymbol{\varphi})$ in $\boldsymbol{L}^2(\Omega_T)$, it follows that

$$-\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}_{hk}^-(t), \nabla \times [\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t \to -\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}(t), \nabla \times [\boldsymbol{m}(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t.$$

By (2.12), it holds that

$$-\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times \boldsymbol{m}(t), \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t - \frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}(t), \nabla \times [\boldsymbol{m}(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t$$

$$= -\ell_{\mathrm{dm}} \int_0^T \langle \nabla \times \boldsymbol{m}(t), \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t - \frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle\!\langle \gamma_T[\boldsymbol{m}(t)], \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle\!\rangle \, \mathrm{d}t,$$

which proves (2.13) for any smooth test function $\boldsymbol{\varphi}$. The desired result then follows by density.

The energy inequality (2.14) is obtained by passing (2.36) to the limit as $h, k \to 0$ and using the available convergence results (2.48), assumption (2.28) on the initial condition, the fact that $k/h \to 0$, in combination with standard lower semicontinuity arguments.

- **Step 2:** Proof of the result for PF-TPS1.

The proof follows the lines of the one for TPS1 discussed in Step 1. In the proof of the variational formulation (2.13), the only difference is the convergence of the second term on the left-hand side of (2.51), which is more subtle here, since omitting the nodal projection the uniform boundedness of $\boldsymbol{m}_{hk}^-$ in $\boldsymbol{L}^\infty(\Omega_T)$ is lost. To show the desired convergence, we start with recalling the so-called Lagrange identity

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \quad \text{for all } \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^3 \tag{2.52}$$

and the continuous embedding $\boldsymbol{H}^s(\Omega) \subset \boldsymbol{L}^4(\Omega)$, which holds for all $s \geq 3/4$. Choosing an

arbitrary $3/4 \leq s < 1$, we obtain the estimate

$$
\begin{aligned}
\left\| |\boldsymbol{m}_{hk}^-|^2 - 1 \right\|_{L^2(\Omega_T)}^2 &= \left\| |\boldsymbol{m}_{hk}^-|^2 - |\boldsymbol{m}|^2 \right\|_{L^2(\Omega_T)}^2 = \int_0^T \left\| |\boldsymbol{m}_{hk}^-(t)|^2 - |\boldsymbol{m}(t)|^2 \right\|_{L^2(\Omega)}^2 \mathrm{d}t \\
&= \int_0^T \left\| [\boldsymbol{m}_{hk}^-(t) + \boldsymbol{m}(t)] \cdot [\boldsymbol{m}_{hk}^-(t) - \boldsymbol{m}(t)] \right\|_{L^2(\Omega)}^2 \mathrm{d}t \\
&\leq \int_0^T \|\boldsymbol{m}_{hk}^-(t) + \boldsymbol{m}(t)\|_{\boldsymbol{L}^4(\Omega)}^2 \|\boldsymbol{m}_{hk}^-(t) - \boldsymbol{m}(t)\|_{\boldsymbol{L}^4(\Omega)}^2 \mathrm{d}t \\
&\leq \int_0^T \|\boldsymbol{m}_{hk}^-(t) + \boldsymbol{m}(t)\|_{\boldsymbol{H}^1(\Omega)}^2 \|\boldsymbol{m}_{hk}^-(t) - \boldsymbol{m}(t)\|_{\boldsymbol{H}^s(\Omega)}^2 \mathrm{d}t \\
&\leq \|\boldsymbol{m}_{hk}^- + \boldsymbol{m}\|_{L^\infty(0,T;\boldsymbol{H}^1(\Omega))}^2 \|\boldsymbol{m}_{hk}^- - \boldsymbol{m}\|_{L^2(0,T;\boldsymbol{H}^s(\Omega))}^2 \\
&\lesssim \|\boldsymbol{m}_{hk}^- - \boldsymbol{m}\|_{L^2(0,T;\boldsymbol{H}^s(\Omega))}^2.
\end{aligned}
$$

Thanks to Lemma 2.5.7, we deduce that $|\boldsymbol{m}_{hk}^-|^2 \to 1$ in $L^2(\Omega_T)$ as $h, k \to 0$. Together with the weak convergence $\boldsymbol{v}_{hk}^- \cdot \boldsymbol{\varphi} \rightharpoonup \partial_t \boldsymbol{m} \cdot \boldsymbol{\varphi}$ in $L^2(\Omega_T)$, it follows that

$$
\begin{aligned}
\int_0^T &\langle \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \\
&\stackrel{(2.52)}{=} \int_0^T \langle |\boldsymbol{m}_{hk}^-(t)|^2, \boldsymbol{v}_{hk}^-(t) \cdot \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \to \int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t.
\end{aligned}
$$

Finally, passing the discrete energy law (2.41) to the limit as $h, k \to 0$, thanks to (2.28), (2.47), the available convergence results (2.48), and standard lower semicontinuity arguments, we obtain (2.14).

- **Step 3:** Proof of the result for TPS2.

The verification of the variational formulation (2.13) follows by the same method used in Step 1 for TPS1. Given an arbitrary $\boldsymbol{\varphi} \in \boldsymbol{C}^\infty(\overline{\Omega_T})$, for any $0 \leq i \leq N-1$ and $t \in (t_i, t_{i+1})$, we test (2.26) with $\boldsymbol{\phi}_h = \boldsymbol{\mathcal{I}}_h[\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \in \boldsymbol{\mathcal{K}}_h(\boldsymbol{m}_h^i)$ to obtain

$$
\begin{aligned}
\alpha \int_0^T &\langle W_{M(k)}(\lambda_{hk}^-(t)) \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \\
&+ \int_0^T \langle \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \\
&+ \ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{\nabla}[\boldsymbol{m}_{hk}^-(t) + (1 + \rho(k))(k/2)\boldsymbol{v}_{hk}^-(t)], \boldsymbol{\nabla}[\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t + \mathcal{O}(h) \qquad (2.53) \\
&= -\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times [\boldsymbol{m}_{hk}^-(t) + (k/2)\boldsymbol{v}_{hk}^-(t)], \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \\
&\quad - \frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}_{hk}^-(t) + (k/2)\boldsymbol{v}_{hk}^-(t), \nabla \times [\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t,
\end{aligned}
$$

where, in analogy with (2.27), we define the piecewise time reconstruction λ_{hk}^- by $\lambda_{hk}^-(t) := \lambda_h^i$ for all $0 \leq i \leq N-1$ and $t \in [t_i, t_{i+1})$.

With the available convergence result (2.48) and the convergence properties of $W_{M(k)}(\cdot)$ and $\rho(\cdot)$, each of the three terms on the left-hand side converges towards the corresponding term of (2.13) as $h, k \to 0$; see [AKST14] for details. We discuss the convergence of the two terms on the right-hand side. Since $\nabla \times \boldsymbol{m}_{hk}^- \rightharpoonup \nabla \times \boldsymbol{m}$ and $\boldsymbol{m}_{hk}^- \times \boldsymbol{\varphi} \to \boldsymbol{m} \times \boldsymbol{\varphi}$ in $\boldsymbol{L}^2(\Omega_T)$, it holds that

$$-\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times \boldsymbol{m}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \to -\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times \boldsymbol{m}(t), \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t.$$

Moreover, it holds that

$$-\frac{\ell_{\mathrm{dm}}}{4} k \int_0^T \langle \nabla \times \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \to 0,$$

which follows from (2.47), since

$$\left| k \int_0^T \langle \nabla \times \boldsymbol{v}_{hk}^-(t), \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \right| \lesssim k \|\boldsymbol{\nabla} \boldsymbol{v}_{hk}^-\|_{\boldsymbol{L}^2(\Omega_T)} \|\boldsymbol{m}_{hk}^-\|_{\boldsymbol{L}^\infty(\Omega_T)} \|\boldsymbol{\varphi}\|_{\boldsymbol{L}^2(\Omega_T)}$$

$$\lesssim k \|\boldsymbol{\nabla} \boldsymbol{v}_{hk}^-\|_{\boldsymbol{L}^2(\Omega_T)}.$$

Hence, the first term on the right-hand side of (2.53) converges towards

$$-\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \nabla \times \boldsymbol{m}(t), \boldsymbol{m}(t) \times \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t.$$

Similarly, we show that the second term on the right-hand side converges towards

$$-\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}(t), \nabla \times [\boldsymbol{m}(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t.$$

As shown in Step 1 for TPS1, it holds that

$$-\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}_{hk}^-(t), \nabla \times [\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t \to -\frac{\ell_{\mathrm{dm}}}{2} \int_0^T \langle \boldsymbol{m}(t), \nabla \times [\boldsymbol{m}(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t.$$

On the other hand, we have that

$$\left| k \int_0^T \langle \boldsymbol{v}_{hk}^-(t), \nabla \times [\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t \right| \lesssim k \|\boldsymbol{v}_{hk}^-\|_{\boldsymbol{L}^2(\Omega_T)} \|\boldsymbol{m}_{hk}^-\|_{\boldsymbol{H}^1(\Omega_T)} \|\boldsymbol{\varphi}\|_{\boldsymbol{W}^{1,\infty}(\Omega_T)} \lesssim k,$$

which shows that

$$-\frac{\ell_{\mathrm{dm}}}{4} k \int_0^T \langle \boldsymbol{v}_{hk}^-(t), \nabla \times [\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}(t)] \rangle \, \mathrm{d}t \to 0.$$

This proves (2.13) for any smooth test function $\boldsymbol{\varphi}$. By density, we obtain the desired result.

Finally, the energy inequality (2.14) is obtained by passing to the limit as $h, k \to 0$ the discrete energy law (2.44) and using standard lower semicontinuity arguments. $\qquad \square$

3 The mass-lumped midpoint scheme for computational micromagnetics: Newton linearization and application to magnetic skyrmion dynamics

This chapter consists of an ongoing work together with Giovanni Di Fratta, Dirk Praetorius, and Michele Ruggeri.

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.

3.1 Introduction

3.1.1 Energetics of a ferromagnet

In the continuum theory of micromagnetism, whose origin dates back to the seminal work of Landau–Lifshitz [LL35] 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* $\boldsymbol{M} \colon \Omega \to \mathbb{R}^3$. Its module, $M_\mathrm{s} := |\boldsymbol{M}|$, describes the so-called *saturation magnetization*. In single-crystal ferromagnets [AFM06, AD15], M_s depends only on the temperature and is assumed to be constant when the specimen is well below the so-called Curie temperature of the material. In this case, the magnetization can be represented in the form $\boldsymbol{M} := M_\mathrm{s}\boldsymbol{m}$, 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

64

minimize the micromagnetic energy functional [Bro63, HS98]

$$\mathcal{E}(\boldsymbol{m}) := \mathcal{E}_\Omega(\boldsymbol{m}) + \mathcal{K}_\Omega(\boldsymbol{m}) + \mathcal{W}_\Omega(\boldsymbol{m}) + \mathcal{A}_\Omega(\boldsymbol{m}) + \mathcal{Z}_\Omega(\boldsymbol{m}) \tag{3.1}$$

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

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)* [Dzy58, Mor60], and accounts for antisymmetric exchange interactions caused by possible lacks 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 [TRJF12, SCR$^+$13]. The third term, $\mathcal{W}_\Omega(\boldsymbol{m})$, is the *magnetostatic self-energy*, i.e., the energy due to the stray field $\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m})$ induced by $M_{\mathrm{s}}\boldsymbol{m}$. From the mathematical point of view, $\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{m})$ can be characterized as the projection of $(-M_{\mathrm{s}}\boldsymbol{m}) \in L^2(\mathbb{R}^3; \mathbb{R}^3)$ on the closed subspace of gradient vector fields $\nabla H^1(\mathbb{R}^3; \mathbb{R}) := \{\nabla u \colon u \in H^1(\mathbb{R}^3; \mathbb{R})\}$ (see, e.g., [Pra04, DMRS20])[1]. Here, μ_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 $\varphi_{\mathrm{an}} \colon \mathbb{S}^2 \to \mathbb{R}_{\geq 0}$ 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 $\boldsymbol{H}_{\mathrm{ext}} \in L^2(\Omega; \mathbb{R}^3)$, assumed to be unaffected by variations of $\boldsymbol{m}$. The competition among the energy contributions in (3.1) explains *most* of the striking spin textures observable in ferromagnetic materials [HS98], in particular, the emergence of magnetic skyrmions [FCS13, FRC17].

3.1.2 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 Ω occupied by the ferromagnet, and one has to model spin interactions among different magnetic materials at their touching interface [AFM06]. The easiest way is to assume a *strong coupling condition* [AD15, ADMN21, DD20]: 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 (3.1), but with the quantities $A = A(\boldsymbol{x})$, $D = D(\boldsymbol{x})$, and $M_{\mathrm{s}} = M_{\mathrm{s}}(\boldsymbol{x})$ to be understood as functions defined on Ω.

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

---

[1]Here, with a slight abuse of notation, we identify $\boldsymbol{m}$ with its extension by zero to the whole $\mathbb{R}^3$.

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

$$g_{\mathrm{asym}}(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{\xi}) = \sum_{d=1}^{3} \boldsymbol{K}_d(\boldsymbol{x})\boldsymbol{\xi}_d \cdot \boldsymbol{s} \quad \text{for all } \boldsymbol{x} \in \Omega, \ \boldsymbol{s} \in \mathbb{R}^3, \text{ and } \boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) \in \mathbb{R}^{3\times3},$$

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

$$g_{\mathrm{sym}}(\boldsymbol{x}, \boldsymbol{\xi}) = \frac{1}{2} \sum_{d=1}^{3} \boldsymbol{A}_d(\boldsymbol{x})\boldsymbol{\xi}_d \cdot \boldsymbol{\xi}_d \quad \text{for all } \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) \in \mathbb{R}^{3\times3},$$

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

$$
\begin{aligned}
g(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{\xi}) &= \frac{1}{2} \sum_{d=1}^{3} \left(\boldsymbol{A}_d(\boldsymbol{x})\boldsymbol{\xi}_d \cdot \boldsymbol{\xi}_d - 2\boldsymbol{K}_d(\boldsymbol{x})\boldsymbol{s} \cdot \boldsymbol{\xi}_d \right) \\
&= \frac{1}{2} \sum_{d=1}^{3} \boldsymbol{A}_d(\boldsymbol{x}) \left(\boldsymbol{\xi}_d - \boldsymbol{A}_d^{-1}(\boldsymbol{x})\boldsymbol{K}_d(\boldsymbol{x})\boldsymbol{s} \right) \cdot \left(\boldsymbol{\xi}_d - \boldsymbol{A}_d^{-1}(\boldsymbol{x})\boldsymbol{K}_d(\boldsymbol{x})\boldsymbol{s} \right) \qquad (3.2) \\
&\quad + \frac{1}{2} \sum_{d=1}^{3} \boldsymbol{K}_d(\boldsymbol{x})\boldsymbol{A}_d^{-1}(\boldsymbol{x})\boldsymbol{K}_d(\boldsymbol{x})\boldsymbol{s} \cdot \boldsymbol{s}.
\end{aligned}
$$

Note that $\boldsymbol{K}_d(\boldsymbol{x})\boldsymbol{A}_d^{-1}(\boldsymbol{x})\boldsymbol{K}_d(\boldsymbol{x})$ is a symmetric matrix. This discussion suggests the opportunity to investigate an energy functional covering the above generalized form; see (3.4) 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 [ADMN21, DD20].

3.1.3 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) [LL35, Gil55], which in the so-called Gilbert form reads

$$\partial_t \boldsymbol{m} = -\gamma_0\, \boldsymbol{m} \times \boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m}) + \alpha\, \boldsymbol{m} \times \partial_t \boldsymbol{m}. \qquad (3.3)$$

This phenomenological equation describes the magnetization dynamics as a dissipative precession driven by the effective field $\boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m}) := -\mu_0^{-1} M_{\mathrm{s}}^{-1} \frac{\delta \mathcal{E}(\boldsymbol{m})}{\delta \boldsymbol{m}}$, 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., [Pro01, AJ06, KP06, BP06, GC07, Alo08, BKP08, Cim08, Cim09, AKT12, AKST14, BFF⁺14, AHP⁺14, FT17a, KW18, HPP⁺19, DPP⁺20, AFKL21].

In this work, we consider the mass-lumped midpoint scheme proposed in [BP06]. 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 [BP06] proves unconditional convergence of the finite element approximation towards a weak solution of LLG in the sense of [AS92] 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 [Bar06, Cim09] incorporate the inexact solution of the nonlinear system into the convergence result. On the other hand, the work [PRS18b] focuses on the design and the analysis of effective approaches to treat the nonlocal field contributions.

3.1.4 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 3.1.2. In particular, the present analysis covers DMI, which is not covered by the analysis in [BP06, Bar06, Cim09, PRS18b]. 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 [Bar06] 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, we accurately test the energy conservation properties of the mass-lumped midpoint scheme and extensively compare it with the tangent plane schemes from [Alo08, AKST14, HPP⁺19, DPP⁺20].

3.1.5 Outline

The remainder of the work is organized as follows: We conclude this section by collecting the notation used throughout the paper. In Section 3.2, we describe the mathematical problem under consideration. In Section 3.3, we present the proposed algorithms and state their stability and convergence results. Section 3.4 is devoted to numerical experiments. Finally, in Sections 3.5–3.6, we collect the proofs of the results stated in Section 3.3.

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 $L^2(\Omega; \mathbb{R}^{3 \times 3})$ 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.

3.2 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

$$\mathcal{E}(\boldsymbol{m}) = \frac{1}{2}\,a(\boldsymbol{m}, \boldsymbol{m}) - \langle \boldsymbol{f}, \boldsymbol{m} \rangle_\Omega, \tag{3.4}$$

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

$$a(\boldsymbol{\psi}, \boldsymbol{\phi}) = \sum_{d=1}^{3} \langle \boldsymbol{A}_d(\partial_d \boldsymbol{\psi} - \boldsymbol{J}_d \boldsymbol{\psi}), \partial_d \boldsymbol{\phi} - \boldsymbol{J}_d \boldsymbol{\phi} \rangle_\Omega - \langle \boldsymbol{\pi}(\boldsymbol{\psi}), \boldsymbol{\phi} \rangle_\Omega. \tag{3.5}$$

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 (3.4) covers the extensions of the classical micromagnetic functional discussed in Section 3.1.2; cf. the expression in (3.2).

The existence of minimizers of (3.4) 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

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

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

$$-\langle \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}), \boldsymbol{\phi} \rangle = \left\langle \frac{\delta \mathcal{E}(\boldsymbol{m})}{\delta \boldsymbol{m}}, \boldsymbol{\phi} \right\rangle = \lim_{\delta \to 0} \frac{\mathcal{E}(\boldsymbol{m} + \delta \boldsymbol{\phi}) - \mathcal{E}(\boldsymbol{m})}{\delta} \overset{(3.4)}{=} a(\boldsymbol{m}, \boldsymbol{\phi}) - \langle \boldsymbol{f}, \boldsymbol{\phi} \rangle_\Omega. \quad (3.6)$$

Turning to the dynamical case, a non-equilibrium configuration $\boldsymbol{m}(t) \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ evolves according to (3.3), which, after a suitable rescaling, reads

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) + \alpha\, \boldsymbol{m} \times \partial_t \boldsymbol{m}, \tag{3.7}$$

with $\alpha > 0$ being the Gilbert damping parameter. The dynamics is dissipative in the sense that any sufficiently smooth solution of (3.7) satisfies the energy law

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(\boldsymbol{m}(t)) = -\alpha \|\partial_t \boldsymbol{m}(t)\|_{\boldsymbol{L}^2(\Omega)}^2 \leq 0 \quad \text{for all } t > 0. \tag{3.8}$$

We conclude this section by recalling the notion of a weak solution of (3.7); see [AS92].

Definition 3.2.1. *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* (3.7) *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

$$
\int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle_\Omega \, \mathrm{d}t
$$
$$
= -\int_0^T \langle \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}(t)), \boldsymbol{\varphi}(t) \times \boldsymbol{m}(t) \rangle \, \mathrm{d}t + \alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle_\Omega \, \mathrm{d}t; \tag{3.9}
$$

(iv) It holds that

$$
\mathcal{E}(\boldsymbol{m}(T)) + \alpha \int_0^T \| \partial_t \boldsymbol{m}(t) \|_{\boldsymbol{L}^2(\Omega)}^2 \mathrm{d}t \le \mathcal{E}(\boldsymbol{m}^0). \tag{3.10}
$$

We note that (3.6) implicitly includes natural boundary conditions on $\boldsymbol{m}$, which are homogeneous Neumann boundary conditions $\partial_{\boldsymbol{n}} = \boldsymbol{0}$ if $\boldsymbol{A}_d = \ell_{\mathrm{ex}}^2 \mathbf{Id}$ and $\boldsymbol{J}_d = \boldsymbol{0}$ for $d = 1, 2, 3$. For a more explicit presentation, we refer to [HPP$^+$19]. The variational formulation (3.9) comes from a weak formulation of (3.7) in the space-time cylinder. The energy inequality (3.10) is a weak counterpart of the dissipative energy law (3.8).

Remark 3.2.2. *(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 (3.8) and the weak form (3.10) of the energy law of LLG read*

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

and

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

respectively, where $\widetilde{\mathcal{E}}(\boldsymbol{m}) = \mathcal{E}(\boldsymbol{m}) + \langle \boldsymbol{f}, \boldsymbol{m} \rangle_\Omega = a(\boldsymbol{m}, \boldsymbol{m})/2$.
(ii) The present setting covers and generalizes the model problems considered in previous mathematical works on the numerical integration of LLG.

- *With the choices $\boldsymbol{A}_d = \ell_{\mathrm{ex}}^2 \mathbf{Id}$ and $\boldsymbol{J}_d = \boldsymbol{0}$ for $d = 1, 2, 3$, where $\ell_{\mathrm{ex}} > 0$ 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*

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

This is the so-called small particle limit of LLG, for which finite element schemes have been proposed, e.g., in the seminal papers [BP06, Alo08].

- *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*

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

  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 [AKT12, BFF⁺14, AKST14, PRS18b, DPP⁺20].

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

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

  which is the setting analyzed in [HPP⁺19] for the simulation of chiral magnetic skyrmions by the means of a family of tangent plane integrators.

3.3 Numerical algorithms and main results

3.3.1 Preliminaries

Let $\kappa \geq 1$. For the spatial discretization, assuming Ω to be a polyhedral domain, we consider a κ-quasi-uniform family $\{\mathcal{T}_h\}_{h>0}$ of regular tetrahedral triangulations of Ω parametrized by the mesh size $h = \max_{K \in \mathcal{T}_h} \text{diam}(K)$, i.e., $\kappa^{-1} h \leq \text{vol}(K)^{1/3} \leq h$ 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

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

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 (3.7) are sought in the *set of admissible approximate magnetizations*

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

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

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

---

[2]Here, $[\boldsymbol{e}_d]_\times$ denotes the 3-by-3 matrix such that $[\boldsymbol{e}_d]_\times \boldsymbol{u} = \boldsymbol{e}_d \times \boldsymbol{u}$ for all $\boldsymbol{u} \in \mathbb{R}^3$.

Using the definition of the nodal interpolant, we see that

$$\langle \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_h = \sum_{\boldsymbol{z} \in \mathcal{N}_h} \beta_{\boldsymbol{z}} \, \boldsymbol{\psi}(\boldsymbol{z}) \cdot \boldsymbol{\phi}(\boldsymbol{z}), \quad \text{where } \beta_{\boldsymbol{z}} := \int_\Omega \varphi_{\boldsymbol{z}} \, \mathrm{d}\boldsymbol{x} > 0. \tag{3.11}$$

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

$$\|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \leq \|\boldsymbol{\phi}_h\|_h \leq \sqrt{5} \, \|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{\phi}_h \in \boldsymbol{V}_h; \tag{3.12}$$

see [Bar15, Lemma 3.9]. Finally, we define the mapping $\mathbb{P}_h \colon \boldsymbol{H}^1(\Omega)^\star \to \boldsymbol{V}_h$ by

$$\langle \mathbb{P}_h \boldsymbol{u}, \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{u}, \boldsymbol{\phi}_h \rangle \quad \text{for all } \boldsymbol{u} \in \boldsymbol{H}^1(\Omega)^\star \text{ and } \boldsymbol{\phi}_h \in \boldsymbol{V}_h, \tag{3.13}$$

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

$$\boldsymbol{\phi}_h^{i+1/2} := \frac{\boldsymbol{\phi}_h^{i+1} + \boldsymbol{\phi}_h^i}{2} \quad \text{and} \quad d_t \boldsymbol{\phi}_h^{i+1} := \frac{\boldsymbol{\phi}_h^{i+1} - \boldsymbol{\phi}_h^i}{k} \quad \text{for all } i \in \mathbb{N}_0, \tag{3.14}$$

as well as the piecewise linear time reconstruction

$$\boldsymbol{\phi}_{hk}(t) := \frac{t - t_i}{k} \boldsymbol{\phi}_h^{i+1} + \frac{t_{i+1} - t}{k} \boldsymbol{\phi}_h^i \quad \text{for all } i \in \mathbb{N}_0 \text{ and } t \in [t_i, t_{i+1}], \tag{3.15}$$

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

3.3.2 Ideal midpoint scheme

In the following algorithm, we adapt the scheme initially proposed in [BP06] 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 (3.13) 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 3.3.1 (ideal midpoint scheme). *Input:* $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$.
Loop: For all $i \in \mathbb{N}_0$, compute $\boldsymbol{m}_h^{i+1} \in \boldsymbol{V}_h$ such that

$$\langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\phi}_h \rangle_h = -\langle \boldsymbol{m}_h^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_h^{i+1/2}), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\phi}_h \rangle_h \tag{3.16}$$

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 3.3.1, we define the piecewise linear time reconstruction $\boldsymbol{m}_{hk}$ via (3.15). In the following theorem, we establish the stability and convergence of the approximations obtained with Algorithm 3.3.1. The proof is postponed to Section 3.5.

Theorem 3.3.2. *(i) Suppose that $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$. Then, for all $i \in \mathbb{N}_0$, (3.16) admits a solution $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$. 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

$$\mathcal{E}(\boldsymbol{m}_h^J) + \alpha k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_h^i\|_h^2 = \mathcal{E}(\boldsymbol{m}_h^0). \tag{3.17}$$

(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 (3.7) in the sense of Definition 3.2.1 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 3.3.3. *Note that, differently from the corresponding estimates for tangent plane schemes [Alo08, AKST14, HPP$^+$19, DPP$^+$20], the stability result for Algorithm 3.3.1 (Theorem 3.3.2(ii)) does not require any geometric assumption on the mesh. Moreover, (3.17) holds with equality and without any artificial dissipative term on the left-hand side.*

Theorem 3.3.2(i) establishes unconditional existence of a solution of (3.16), 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 3.3.1 is well-posed.

Proposition 3.3.4. *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 (3.16) admits a unique solution $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ for all $i \in \mathbb{N}_0$.*

The proof of Proposition 3.3.4 can be obtained simplifying the argument of the proofs of Proposition 3.3.5 and Theorem 3.3.7 below, therefore we omit it.

3.3.3 Practical midpoint schemes

Each time-step of Algorithm 3.3.1 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 (3.16). The second term on the right-hand side, at first glance also nonlinear in $\boldsymbol{m}_h^{i+1}$, turns out to be actually linear. Indeed, it holds that

$$\boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1} \overset{(3.14)}{=} \frac{\boldsymbol{m}_h^{i+1} + \boldsymbol{m}_h^i}{2} \times \frac{\boldsymbol{m}_h^{i+1} - \boldsymbol{m}_h^i}{k} = -\frac{1}{k} \boldsymbol{m}_h^{i+1} \times \boldsymbol{m}_h^i.$$

However, using an arbitrary off-the-shelf nonlinear solver for (3.16), the conservation and stability properties of Algorithm 3.3.1 established in Theorem 3.3.2(i)–(ii) are in general lost. Moreover, $\boldsymbol{\pi}$ can be nonlocal and non-exactly computable (e.g., for the stray field), so that the field contribution $\boldsymbol{\pi}(\boldsymbol{m}_h^{i+1/2})$ must be numerically approximated. Hence, a direct

implementation of Algorithm 3.3.1 should be based on an inner iteration performing the solution of the nonlinear system (3.16) and the approximate computation of $\boldsymbol{\pi}(\boldsymbol{m}_h^{i+1/2})$.

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 (3.16), from which we will obtain two *practical midpoint schemes.*

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

$$a^{\mathrm{loc}}(\boldsymbol{\psi},\boldsymbol{\phi}) = \sum_{d=1}^{3}\langle \boldsymbol{A}_d(\partial_d\boldsymbol{\psi} - \boldsymbol{J}_d\boldsymbol{\psi}), \partial_d\boldsymbol{\phi} - \boldsymbol{J}_d\boldsymbol{\phi}\rangle_\Omega \quad \text{for all } \boldsymbol{\psi},\boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega). \tag{3.18}$$

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

$$\mathcal{E}^{\mathrm{loc}}(\boldsymbol{m}) := \frac{1}{2}\,a^{\mathrm{loc}}(\boldsymbol{m},\boldsymbol{m}) - \langle \boldsymbol{f},\boldsymbol{m}\rangle_\Omega \overset{(3.4)}{=} \mathcal{E}(\boldsymbol{m}) + \frac{1}{2}\langle \boldsymbol{\pi}(\boldsymbol{m}),\boldsymbol{m}\rangle_\Omega$$

and

$$-\langle \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{m}),\boldsymbol{\phi}\rangle := \left\langle \frac{\delta\mathcal{E}^{\mathrm{loc}}(\boldsymbol{m})}{\delta\boldsymbol{m}},\boldsymbol{\phi}\right\rangle \overset{(3.18)}{=} a^{\mathrm{loc}}(\boldsymbol{m},\boldsymbol{\phi}) - \langle \boldsymbol{f},\boldsymbol{\phi}\rangle_\Omega$$

$$\overset{(3.5)}{=} a(\boldsymbol{m},\boldsymbol{\phi}) + \langle \boldsymbol{\pi}(\boldsymbol{m}),\boldsymbol{\phi}\rangle_\Omega - \langle \boldsymbol{f},\boldsymbol{\phi}\rangle_\Omega$$

$$\overset{(3.6)}{=} -\langle \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}),\boldsymbol{\phi}\rangle + \langle \boldsymbol{\pi}(\boldsymbol{m}),\boldsymbol{\phi}\rangle_\Omega,$$

respectively. Then, for $i \in \mathbb{N}_0$, we rewrite (3.16) in terms of the new unknown $\boldsymbol{\eta}_h^i := \boldsymbol{m}_h^{i+1/2} \in \boldsymbol{V}_h$. Since $d_t\boldsymbol{m}_h^{i+1} = 2(\boldsymbol{\eta}_h^i - \boldsymbol{m}_h^i)/k$, it is easy to see that (3.16) is equivalent to the following problem: First, compute $\boldsymbol{\eta}_h^i \in \boldsymbol{V}_h$ such that, for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$, it holds that

$$\langle \boldsymbol{\eta}_h^i,\boldsymbol{\phi}_h\rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^i \times \mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^i),\boldsymbol{\phi}_h\rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^i \times \mathbb{P}_h\boldsymbol{\pi}(\boldsymbol{\eta}_h^i),\boldsymbol{\phi}_h\rangle_h + \alpha\langle \boldsymbol{\eta}_h^i \times \boldsymbol{m}_h^i,\boldsymbol{\phi}_h\rangle_h \tag{3.19}$$
$$= \langle \boldsymbol{m}_h^i,\boldsymbol{\phi}_h\rangle_h.$$

Then, define

$$\boldsymbol{m}_h^{i+1} := 2\,\boldsymbol{\eta}_h^i - \boldsymbol{m}_h^i. \tag{3.20}$$

To treat the nonlocal contribution $\boldsymbol{\pi}(\boldsymbol{\eta}_h^i)$, we adopt the implicit-explicit (IMEX) approach introduced in [PRS18b]. Let $\boldsymbol{\pi}_h\colon \boldsymbol{V}_h \to \boldsymbol{V}_h$ be an operator approximating $\boldsymbol{\pi}$, assumed to be linear and uniformly bounded in $\boldsymbol{L}^2(\Omega)$ in the sense that $\|\boldsymbol{\pi}_h\|_{L(\boldsymbol{L}^2(\Omega);\boldsymbol{L}^2(\Omega))} \leq C_{\boldsymbol{\pi}}$ for some $C_{\boldsymbol{\pi}} > 0$ independent of h. Moreover, we say that $\boldsymbol{\pi}_h$ is *consistent* with $\boldsymbol{\pi}$, if for all $\boldsymbol{\phi} \in \boldsymbol{L}^2(\Omega)$ and all $(\boldsymbol{\phi}_h)_{h>0} \subset \boldsymbol{V}_h$ with $\boldsymbol{\phi}_h \to \boldsymbol{\phi}$ in $\boldsymbol{L}^2(\Omega)$ as $h \to 0$, it holds that

$$\boldsymbol{\pi}_h(\boldsymbol{\phi}_h) \to \boldsymbol{\pi}(\boldsymbol{\phi}) \quad \text{in} \quad \boldsymbol{L}^2(\Omega) \quad \text{as } h \to 0. \tag{3.21}$$

We define $\boldsymbol{m}_h^{-1} := \boldsymbol{m}_h^0$ and

$$\boldsymbol{\Pi}_h(\boldsymbol{m}_h^i,\boldsymbol{m}_h^{i-1}) := \frac{3}{2}\boldsymbol{\pi}_h(\boldsymbol{m}_h^i) - \frac{1}{2}\boldsymbol{\pi}_h(\boldsymbol{m}_h^{i-1}). \tag{3.22}$$

Then, in (3.19), we replace $\boldsymbol{\pi}(\boldsymbol{\eta}_h^i)$ with its approximation $\boldsymbol{\Pi}_h(\boldsymbol{m}_h^i,\boldsymbol{m}_h^{i-1})$ to obtain

$$\langle \boldsymbol{\eta}_h^i,\boldsymbol{\phi}_h\rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^i \times \mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^i),\boldsymbol{\phi}_h\rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^i \times \mathbb{P}_h\boldsymbol{\Pi}_h(\boldsymbol{m}_h^i,\boldsymbol{m}_h^{i-1}),\boldsymbol{\phi}_h\rangle_h \tag{3.23}$$
$$+ \alpha\langle \boldsymbol{\eta}_h^i \times \boldsymbol{m}_h^i,\boldsymbol{\phi}_h\rangle_h = \langle \boldsymbol{m}_h^i,\boldsymbol{\phi}_h\rangle_h.$$

In particular, the nonlocal contribution, treated explicitly, becomes independent of the unknown $\boldsymbol{\eta}_h^i$. We now discuss two strategies to linearize (3.23) in order to arrive at two practical midpoint schemes. To emphasize the inexact solution of (3.23) 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.

3.3.3.1 Constraint-preserving fixed-point iteration

We solve (3.23) 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

$$
\begin{aligned}
\langle \boldsymbol{\eta}_h^{i,\ell+1}, \boldsymbol{\phi}_h \rangle_h &+ \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}), \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}), \boldsymbol{\phi}_h \rangle_h \\
&+ \alpha \langle \boldsymbol{\eta}_h^{i,\ell+1} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h,
\end{aligned}
\tag{3.24}
$$

until

$$
\| \boldsymbol{\mathcal{I}}_h[\boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}))] \|_h \leq \varepsilon,
\tag{3.25}
$$

where $\boldsymbol{\mathcal{I}}_h[\cdot]$ denotes the vector-valued nodal interpolant. If $\ell^* \in \mathbb{N}_0$ is the smallest integer for which the stopping criterion (3.25) is satisfied, in view of (3.20), 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 3.6.

Proposition 3.3.5. *Let $i \in \mathbb{N}_0$.*
(i) For all $\ell \in \mathbb{N}_0$, the variational problem (3.24) 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

$$
\| \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1} \|_h \leq q \| \boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell} \|_h \quad \text{for all } \ell \in \mathbb{N}_0.
\tag{3.26}
$$

The constants q, h_0, k_0 depend only on the mesh parameter κ and the problem data.
(iii) Under the assumptions of part (ii), the stopping criterion (3.25) is met in a finite number of iterations. If $\ell^* \in \mathbb{N}_0$ denotes the smallest integer for which (3.25) 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 $\boldsymbol{\mathcal{M}}_h$.

For all $i \in \mathbb{N}_0$, let $\boldsymbol{r}_{h\varepsilon}^i := \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell^*+1}) - \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell^*})) \in \boldsymbol{V}_h$. Because of the stopping criterion (3.25), it holds that $\|\boldsymbol{\mathcal{I}}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \leq \varepsilon$. With this definition, the proposed linearization of Algorithm 3.3.1 is covered by the following algorithm.

Algorithm 3.3.6 (practical midpoint scheme, constraint preserving fixed-point iteration)**.**
Input: $\boldsymbol{m}_{h\varepsilon}^0 := \boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$.
Loop: For all $i \in \mathbb{N}_0$, use the constraint preserving fixed-point iteration (3.24)–(3.25) to compute $\boldsymbol{m}_{h\varepsilon}^{i+1} \in \boldsymbol{\mathcal{M}}_h$ and $\boldsymbol{r}_{h\varepsilon}^i \in \boldsymbol{V}_h$ with $\|\boldsymbol{\mathcal{I}}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \leq \varepsilon$ such that

$$
\begin{aligned}
\langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h = &-\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h \\
&+ \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times [\boldsymbol{r}_{h\varepsilon}^i + \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}))], \boldsymbol{\phi}_h \rangle_h
\end{aligned}
\tag{3.27}
$$

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 3.3.6. The proof is postponed to Section 3.6.

Theorem 3.3.7. *(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 $k < k_0$, (3.27) admits solutions $\boldsymbol{m}_{h\varepsilon}^{i+1} \in \mathcal{M}_h$ and $\boldsymbol{r}_{h\varepsilon}^i \in \boldsymbol{V}_h$ with $\|\boldsymbol{\mathcal{I}}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \leq \varepsilon$ 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 κ 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

$$\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^J) + \alpha k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2 = \mathcal{E}(\boldsymbol{m}_h^0) \tag{3.28}$$

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

(iii) Let $T > 0$. Let $J \in \mathbb{N}$ be the smallest integer such that $T \leq kJ$. 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 $0 < h_0^* \leq h_0$, $0 < k_0^* \leq k_0$, and $\varepsilon_0^* > 0$ such that, for all $h < h_0^*$, $k < k_0^*$, $\varepsilon < \varepsilon_0^*$, and $1 \leq j \leq J$, we have the stability estimate

$$\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{i=0}^{j-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{L}^2(\Omega)}^2 \leq C. \tag{3.29}$$

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

(iv) Additionally 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 (3.21) 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, $\boldsymbol{m}_{h\varepsilon k}|_{\Omega_T} \stackrel{\star}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(0, T; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and $\boldsymbol{m}_{h\varepsilon k}|_{\Omega_T} \rightharpoonup \boldsymbol{m}$ in $\boldsymbol{H}^1(\Omega_T)$ as $h, k, \varepsilon \to 0$. The limit function $\boldsymbol{m}$ satisfies the conditions (i)–(iv) of Definition 3.2.1.

3.3.3.2 Newton iteration

Based on the Newton scheme, in [BBNP14, Section 1.4.1] the authors employ a linearization of the nonlinear system (3.16) 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 3.3.3.1.

For three dimensional micromagnetics and considering the full effective field (3.6), in Section 3.7.3 we apply Newton's method to the nonlinear system of equations (3.23) 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 $\boldsymbol{\eta}_h^{i,\ell} \in \boldsymbol{V}_h$, compute $\boldsymbol{u}_h^{i,\ell} \in \boldsymbol{V}_h$ such that, for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$, it holds that

$$
\begin{aligned}
&\langle \boldsymbol{u}_h^{i,\ell}, \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \boldsymbol{u}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}) + \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \boldsymbol{u}_h^{i,\ell} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h \\
&\quad + \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{u}_h^{i,\ell}) - \boldsymbol{f}), \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{m}_{h\varepsilon}^i - \boldsymbol{\eta}_h^{i,\ell}, \boldsymbol{\phi}_h \rangle_h \\
&\qquad - \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}) + \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h \rangle_h - \alpha \langle \boldsymbol{\eta}_h^{i,\ell} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h,
\end{aligned}
\tag{3.30}
$$

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

$$
\| \boldsymbol{\mathcal{I}}_h \big[\boldsymbol{u}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{u}_h^{i,\ell}) - \boldsymbol{f}) \big] \|_h \leq \varepsilon.
\tag{3.31}
$$

If $\ell^* \in \mathbb{N}_0$ is the smallest integer for which the stopping criterion (3.31) 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 $\boldsymbol{r}_{h\varepsilon}^i = \boldsymbol{\mathcal{I}}_h \big[\boldsymbol{u}_h^{i,\ell^*} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{u}_h^{i,\ell^*}) - \boldsymbol{f}) \big] \in \boldsymbol{V}_h$. In view of the stopping criterion (3.31), it holds that $\| \boldsymbol{r}_{h\varepsilon}^i \|_h \leq \varepsilon$. With this definition, the proposed linearization of Algorithm 3.3.1 based on the Newton method is covered by the following algorithm.

Algorithm 3.3.8 (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* (3.30)–(3.31) *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*

$$
\begin{aligned}
\langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h = &-\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h \\
&+ \langle \boldsymbol{r}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h + \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h \rangle_h
\end{aligned}
\tag{3.32}
$$

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

The results on Algorithm 3.3.8 are stated in Lemma 3.3.9 ($\boldsymbol{L}^\infty$-bound), Theorem 3.3.10 (stability), and Theorem 3.3.11 (well-posedness) below. Compared to Algorithm 3.3.6, our analysis is more involved: Precisely, for $i < J$ the proof of Theorem 3.3.11 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 3.3.8, while Lemma 3.3.9 and Theorem 3.3.10 require termination of (3.30)–(3.31) so that $\boldsymbol{m}_{h\varepsilon}^{i+1}$ is well-defined.

In contrast to the fixed-point iteration (3.24)–(3.25), the Newton iteration (3.30)–(3.31) 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 3.3.8, in the following lemma we establish uniform $\boldsymbol{L}^\infty(\Omega)$-boundedness of the approximations obtained with Algorithm 3.3.8.

Lemma 3.3.9. *Suppose* $h, k, \varepsilon > 0$*,* $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ *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* (3.30)–(3.31) *in Algorithm 3.3.8 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 3.3.8 and satisfy* (3.32) *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 exists 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 $\|\boldsymbol{m}_{h\varepsilon}^{n+1}\|_{\boldsymbol{L}^\infty(\Omega)} \le C_\infty$ uniformly for all $0 \le n \le i$. The thresholds h_0, k_0, ε_0 depend only on the mesh parameter κ and the problem data, while the bound $C_\infty > 0$ depends only on κ, $\varepsilon h^{-3/2} \lesssim 1$, 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 3.3.8, in the following theorem we establish the stability and convergence of the approximations obtained with Algorithm 3.3.8. The proof is postponed to Section 3.7.1.

Theorem 3.3.10. *Let $T > 0$ and suppose that $\{\mathcal{T}_h\}_{h>0}$ is a κ-quasi-uniform family of triangulations. Suppose $h, k, \varepsilon > 0$, $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ and let $J \in \mathbb{N}$ be the smallest integer such that $T \le kJ$. Let $0 \le i < J$ and suppose that the Newton iteration (3.30)–(3.31) in Algorithm 3.3.8 terminates for all $0 \le n \le i$, 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 3.3.8 and satisfy (3.32) with $\|\boldsymbol{r}_{h\varepsilon}^n\|_h \le \varepsilon$ for all $0 \le n \le i$.*
(i) Under these assumptions, the scheme satisfies the discrete energy identity

$$\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^{i+1}) + \alpha k \sum_{n=0}^{i} \|d_t \boldsymbol{m}_{h\varepsilon}^{n+1}\|_h^2 = \mathcal{E}(\boldsymbol{m}_h^0) \tag{3.33}$$

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

(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 $0 < \varepsilon_0 \le \alpha$ such that, for all $h < h_0$, $k < k_0$, $\varepsilon < \varepsilon_0$ we have the stability estimate

$$\|\boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{n=0}^{i} \|d_t \boldsymbol{m}_{h\varepsilon}^{n+1}\|_{\boldsymbol{L}^2(\Omega)}^2 \le C. \tag{3.34}$$

The constant $C > 0$ and the thresholds h_0, k_0, ε_0 depend only on the mesh parameter κ, the final time T, and the problem data.
(iii) Additionally 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 (3.21) 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, $\boldsymbol{m}_{h\varepsilon k} \overset{\star}{\rightharpoonup} \boldsymbol{m}$ 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 the conditions (i)–(iv) of Definition 3.2.1.

The following theorem guarantees that under appropriate CFL conditions Algorithm 3.3.8 is well-posed, which is required by Lemma 3.3.9 and Theorem 3.3.10.

Theorem 3.3.11. *Let $T > 0$. Suppose $h, k, \varepsilon > 0$, $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ and let $J \in \mathbb{N}$ be the smallest integer such that $T \le kJ$.*
(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 3.3.8 is well defined, i.e., for all $i = 0, \ldots, J-1$ it provides after finitely many iterations of Newton's method (3.30)–(3.31) solutions $\boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{r}_{h\varepsilon}^i \in \boldsymbol{V}_h$ to (3.32) 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 (3.30)–(3.31) required to solve (3.32) is bounded by $\log_2 \log_2(C_\star k h^{-7/2} \varepsilon^{-1})$. The thresholds h_0, k_0, ε_0 and the constant $C_\star$ depend only on the mesh parameter κ and the problem data.

3.3.3.3 Coupling conditions on practical midpoint schemes

While the ideal midpoint scheme (Algorithm 3.3.1) is unconditionally convergent towards a weak solution of LLG, the analysis of the practical midpoint schemes (Algorithm 3.3.6 and Algorithm 3.3.8) crucially relies on CFL conditions imposed on the discretization parameters $h, k, \varepsilon > 0$. We conclude this section by Table 3.1, giving an overview on the imposed coupling conditions sufficient to establish a rigorous analysis of the practical midpoint schemes.

	Fixed-point linearization Algorithm 3.3.6	Newton linearization Algorithm 3.3.8
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 3.1: Sufficient CFL conditions for the analysis of the practical midpoint schemes of Section 3.3.3.

3.4 Numerical experiments

The goal of this section is threefold: First, in Section 3.4.1 we verify the extension of the midpoint scheme to the DMI contribution and its correct implementation by simulating an experiment on skyrmion dynamics from [SCR$^+$13]. The simulation results with the midpoint scheme are compared to theirs and to simulations with the tangent plane scheme from [HPP$^+$19]. In Section 3.4.2, we introduce a variation of the experiment from [SCR$^+$13] 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 3.4.3. 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 [Pfe, PRS$^+$20].

3.4.1 Stability of isolated skyrmions in nanodisks

To validate the extension of the midpoint scheme incorporating the DMI contribution, we reproduce a numerical experiment from [SCR$^+$13] for both the practical midpoint scheme based on the constraint preserving fixed-point iteration (Algorithm 3.3.6) and the practical midpoint scheme based on Newton's method (Algorithm 3.3.8). There, the relaxed states of a thin nanodisk of diameter 80 nm (aligned with x_1x_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.,

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

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 [ngs] 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 [HPP$^+$19, 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 3.3.7 and Theorem 3.3.11. 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 3.1. 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 $D = 3$ mJ/m^2 to the maximum value of circa 48 nm for $D = 6$ mJ/m^2.

In Figure 3.1, 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 [SCR$^+$13, Figure 1] and [HPP$^+$19, Section 4.2]. This validates both the extension of the midpoint scheme to DMI energy contributions and its implementation in Commics [Pfe, PRS$^+$20].
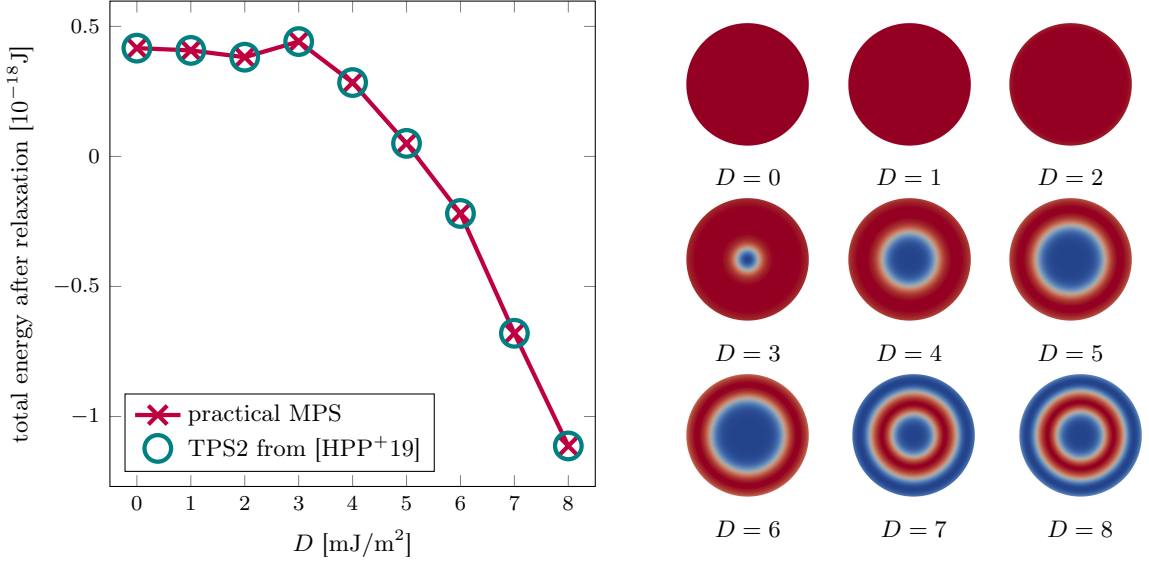
Figure 3.1: Experiment of Section 3.4.1. Left: Final energy for different values of the DMI constant. Right: Magnetization component $\boldsymbol{m}_3$ ranging from -1 (blue) to $+1$ (red) of the relaxed state for different values of the DMI constant (in mJ/m^2). The results computed with either of the practical midpoint schemes, Algorithm 3.3.6 or Algorithm 3.3.8, coincide.

3.4.2 Reliable schemes for energy sensitive dynamics

We recall the discrete energy equality (3.17) 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 [Alo08] we recite from [HPP+19, Proposition 2] the discrete energy inequality

$$\mathcal{E}(\boldsymbol{m}_h^J) + (\alpha - Ck/h)k \sum_{i=0}^{J-1} \|\boldsymbol{v}_h^i\|_h^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2 \sum_{i=0}^{J-1} \|\boldsymbol{\nabla} \boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \le \mathcal{E}(\boldsymbol{m}_h^0)\,, \qquad (3.35)$$

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 (3.35) stems from an inverse estimate used in the analysis of [HPP+19] to control the discrete energy in presence of a DMI energy contribution. The third term on the left-hand side in (3.35) corresponds to artificial damping introduced by implicit treatment in time of the Laplacian for $1/2 < \theta \le 1$, 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 [AKST14], which provides a discrete energy inequality, which, although not

identical to (3.35), 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 [DPP$^+$20]. 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 3.4.1: Considering the different relaxed states in Figure 3.1(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 3.1(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_{crit}^{2-3} and D_{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_{crit}^{2-3} and D_{crit}^{6-7}. For all three schemes, dynamics are simulated with identical parameters:

We consider the fixed mesh from Section 3.4.1. Although this mesh does not satisfy the so-called *angle condition* ensuring validity of (3.35), stability of the tangent plane integrators is still recovered for the smaller time-step sizes meeting $k = o(h^2)$ in this experiment; see [HPP$^+$19, (15) and Remark 3(iv)]. To narrow down the critical values D_{crit}^{2-3} and D_{crit}^{6-7}, we simulate the relaxation dynamics for different values of the DMI constant D corresponding to a resolution of $0.0025\,\mathrm{mJ/m^2}$ as seen in Figure 3.2. For each of the integrators and all considered DMI constants D, we relax the initial state using time-step sizes $k = 1/100\,\mathrm{ps}, 1/200\,\mathrm{ps}, 1/400\,\mathrm{ps}, 1/800\,\mathrm{ps}, 1/1600\,\mathrm{ps}$, 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 $k = 1/100\,\mathrm{ps}, 1/200\,\mathrm{ps}$. 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 3.2 show a sharp transition D_{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 3.3 for a compilation of simulation details on this interval of broken states. While for the tangent plane integrators the determined transition value D_{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 (3.35).
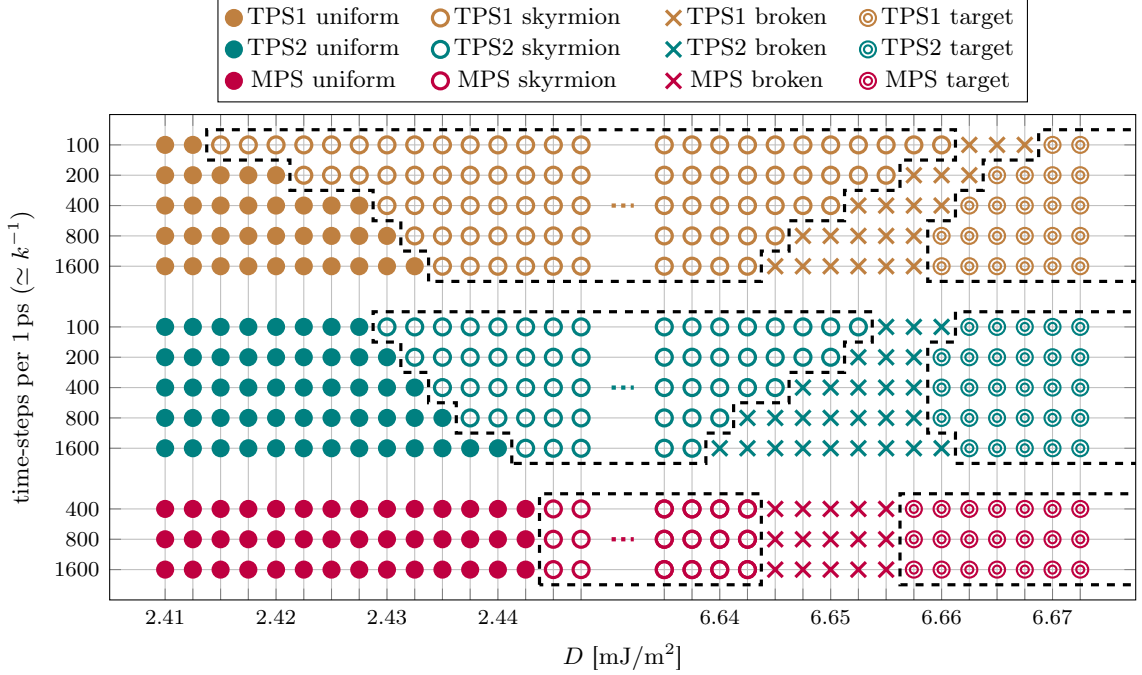
Figure 3.2: Experiment of Section 3.4.2. Each marker corresponds to one simulation carried out with one of the three integrators, for a DMI parameter D with one particular time-step size k. The shape of a marker characterizes the qualitative result after relaxing the skyrmion-like initial state until 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 3.3.6 or Algorithm 3.3.8, coincide.

We conclude that the tangent plane schemes are preferable for uncritical simulations as in Section 3.4.1 or [HPP⁺19, 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.

3.4.3 Numerical study on the CFL conditions

Our results from Section 3.3.3.1 and Section 3.3.3.2, respectively, provide sufficient CFL conditions guaranteeing well-posedness and stability of the practical midpoint schemes in Theorem 3.3.7(i) and Theorem 3.3.11. In this section we investigate whether the CFL
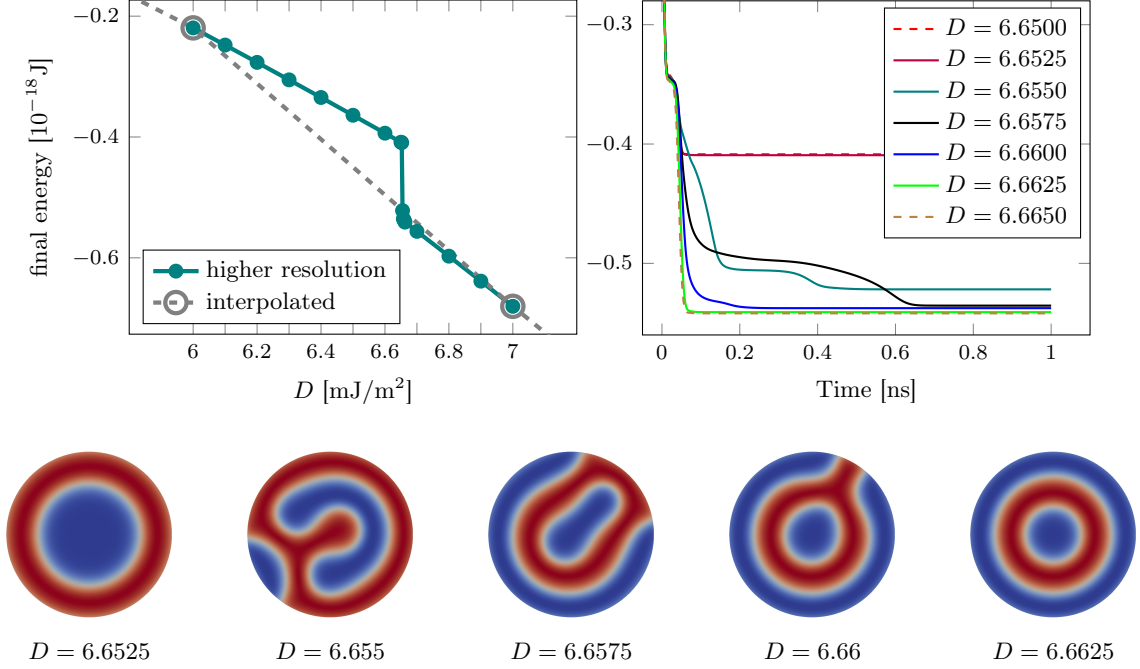
Figure 3.3: Experiment of Section 3.4.2. Details on the simulations for critical transition values between $D = 6\,\mathrm{mJ/m^2}$ and $D = 7\,\mathrm{mJ/m^2}$ with the second-order tangent plane scheme and time-step size $k = 1/100\,\mathrm{ps}$. Left: Critical area of Figure 3.1(left) recomputed with 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 3.2), qualitatively coincides with one of the three broken-symmetry states obtained by TPS2 and $k = 1/100\,\mathrm{ps}$ displayed here.

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 $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta} \boldsymbol{m}$, 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 [PRS$^+$20, Section 5.2], leading to a uniform mesh $\mathcal{T}_h$ of congruent tetrahedra, each of diameter $h_{\max} = \sqrt{3}/N$ and with shortest edge length $h_{\min} = 1/N$. 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 $\boldsymbol{0} \neq \boldsymbol{z} \in \mathcal{N}_h$.

3.4.3.1 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.

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 (3.25) or (3.31), respectively, in the first time-step of Algorithm 3.3.6 or Algorithm 3.3.8. The results depicted in Figure 3.4 show that for both practical midpoint
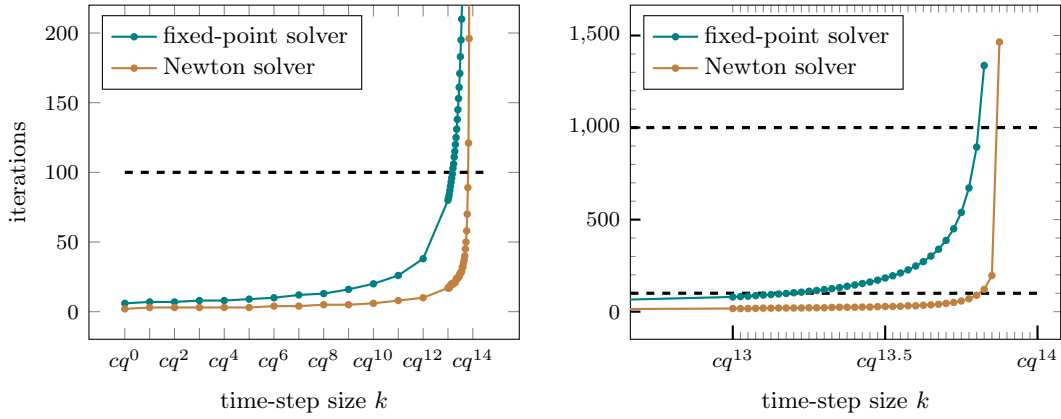
Figure 3.4: Experiment of Section 3.4.3.1. Left: The number of nonlinear iterations rapidly grows as the time-step size $k = cq^j > 0$ with $c = 0.00016$ and $q = 5/4$ approaches the threshold value $k \to k_{\text{thresh}}(h)$. Right: Zoom into the critical area between cq^{13} and cq^{14} where the blow-up occurs.

schemes the number of nonlinear iterations stays well-bounded until a certain threshold value $k_{\text{thresh}}(h) > 0$ 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 $k < k_{\text{thresh}}(h)$ as feasible, and those with $k > k_{\text{thresh}}(h)$ as non-feasible. Surprisingly, despite the different theoretical CFL conditions $k = o(h^2)$ and $k = o(h^{7/3})$ imposed in Proposition 3.3.5 and Theorem 3.3.11, respectively, the threshold value $k_{\text{thresh}}(h)$ seems to coincide for Algorithm 3.3.6 and Algorithm 3.3.8. This observation is investigated further in Section 3.4.3.2. 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 3.4.3.2 we will use the following criterion to classify feasibility of discretization parameters: If for any given (h, k, ε) the respective stopping criterion (3.25) or (3.31), is not met after at most 100 iterations of the nonlinear solver in Algorithm 3.3.6 or Algorithm 3.3.8, 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 $k_{\text{thresh}}(h) > 0$ such that the nonlinear solver converges for $0 < k < k_{\text{thresh}}(h)$ and diverges for $k > k_{\text{thresh}}(h)$. Although only an approximation, Figure 3.4 shows that in view of practical applicability this estimation of $k_{\text{thresh}}(h)$ seems quite appropriate as nonlinear iteration numbers increase drastically as k approaches $k_{\text{thresh}}(h)$.

3.4.3.2 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 3.3.7(i) and Theorem 3.3.11, 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 $k \in \{0.00016 \cdot \left(\frac{5}{4}\right)^j : j = 0, \dots, 27\}$, 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 3.4.3.1, the threshold value of 100 nonlinear iterations is used to classify feasibility of the discretization parameters.
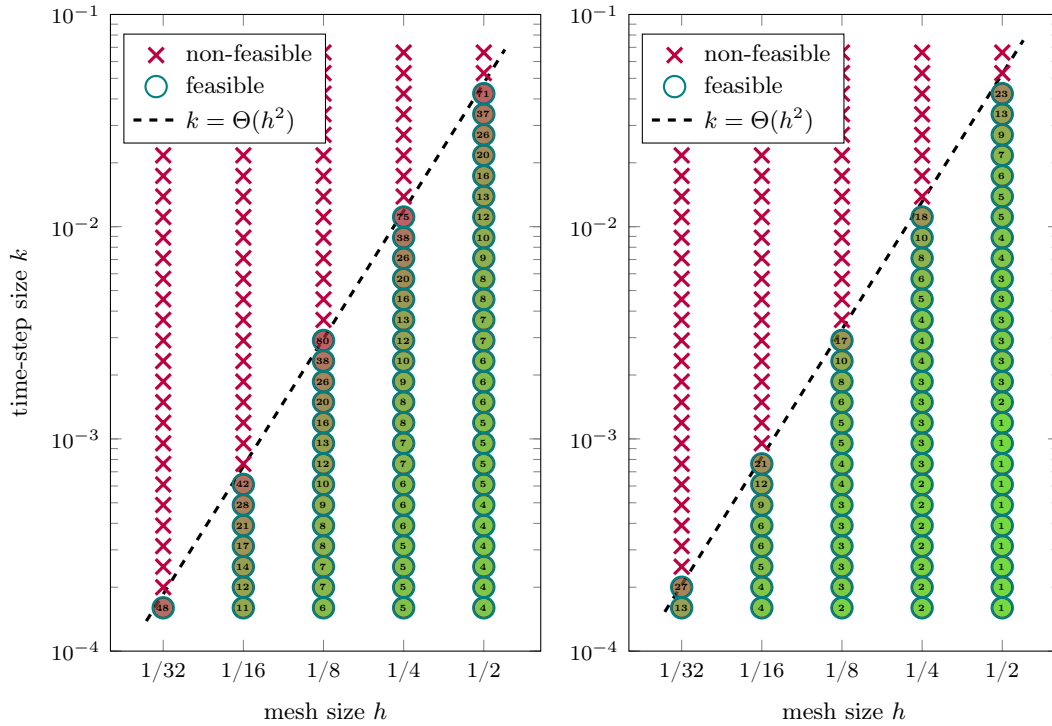
Figure 3.5: Experiment of Section 3.4.3.2. Convergence of the nonlinear solvers in the practical midpoint schemes is investigated. For feasible parameters, the number of nonlinear iterations is given inside the circle. Left: Practical midpoint scheme based on the fixed-point iteration (Algorithm 3.3.6). Right: Practical midpoint scheme based on the Newton iteration (Algorithm 3.3.8). The data points show feasibility if $k = \Theta(h^\beta)$ with possible slopes $1.93 \leq \beta_{\text{fixed-point}} \leq 2.09$ and $1.85 \leq \beta_{\text{newton}} \leq 2.01$.

The results of this experiment shown in Figure 3.5 give insight to the applicability of the practical midpoint schemes: First, for Algorithm 3.3.6 the theoretically sufficient CFL condition $k = o(h^2)$ is shown to be 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 3.3.11 from $k = o(h^{7/3})$ to $k = o(h^2)$. We note that also in the simulation of skyrmion dynamics in Section 3.4.2 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).

3.4.3.3 Constraint violation induced by nonlinear solver accuracy

In contrast to the fixed-point iteration from Section 3.3.3.1, the Newton iteration from Section 3.3.3.2 does not inherently preserve discrete unit-length, i.e., $\boldsymbol{m}_{h\varepsilon}^i \notin \boldsymbol{\mathcal{M}}_h$ 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 $k_{\mathrm{thresh}}(h)$ from Section 3.4.3.1. In Figure 3.6 we plot the deviations

$$\max_{\boldsymbol{z} \in \mathcal{N}_h} |\boldsymbol{m}_{h\varepsilon k}(T, \boldsymbol{z})| - 1 \qquad \text{and} \qquad 1 - \min_{\boldsymbol{z} \in \mathcal{N}_h} |\boldsymbol{m}_{h\varepsilon k}(T, \boldsymbol{z})| \qquad (3.36)$$

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.

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 3.3.7(i) and Theorem 3.3.11, respectively, is yet to be proposed in a future numerical study.

3.5 Proof of Theorem 3.3.2 for the ideal midpoint scheme

3.5.1 Existence of solutions, unit-length constraint, and stability

Proof of Theorem 3.3.2(i). Let $i \in \mathbb{N}_0$ be arbitrary. Define $\boldsymbol{F} \colon \boldsymbol{V}_h \to \boldsymbol{V}_h$ by

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

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 (3.16). Since

$$\langle \boldsymbol{F}(\boldsymbol{\phi}_h), \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{\phi}_h - \boldsymbol{m}_h^i, \boldsymbol{\phi}_h \rangle_h \geq 0 \quad \text{for all } \boldsymbol{\phi}_h \in \boldsymbol{V}_h \text{ with } \|\boldsymbol{\phi}_h\|_h = \|\boldsymbol{m}_h^i\|_h > 0,$$
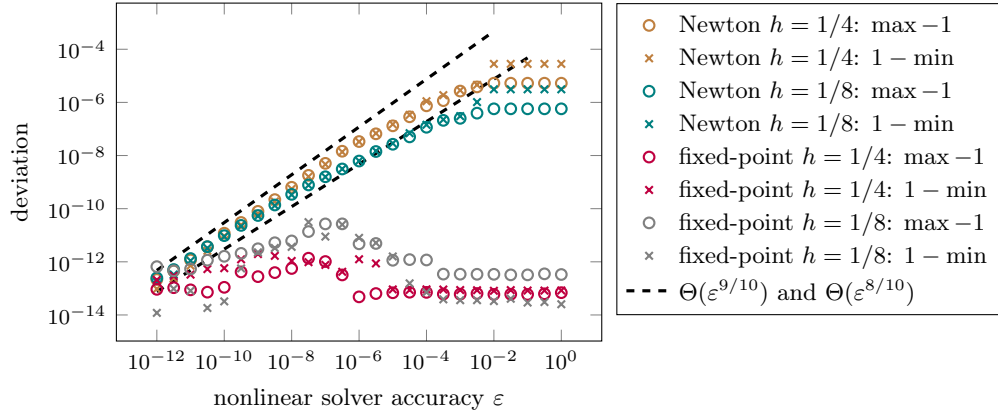
Figure 3.6: Experiment of Section 3.4.3.3. Dependence of the constraint violation (3.36) on the nonlinear solver accuracy $\varepsilon > 0$ is investigated. For the Newton solver the deviation from unit-length decreases as $\varepsilon \to 0$. No obvious correlation is observed for the fixed-point iteration, which is expected, since it is designed to be constraint preserving; see Proposition 3.3.5(iii).

an application of the Brouwer fixed-point theorem (see, e.g., [GR86, 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 (3.16) admits a solution $\boldsymbol{m}_h^{i+1} \in \boldsymbol{V}_h$.

Let $\boldsymbol{z} \in \mathcal{N}_h$ be arbitrary. We test (3.16) with $\boldsymbol{\phi}_h = \boldsymbol{m}_h^{i+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}} \in \boldsymbol{V}_h$ to obtain that

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

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

Proof of Theorem 3.3.2(ii). Let $J \in \mathbb{N}$. To show (3.17), 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 (3.16). 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

$$
\begin{aligned}
\langle \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_h^{i+1/2}), d_t \boldsymbol{m}_h^{i+1} \rangle_h &\overset{(3.13)}{=} \langle \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_h^{i+1/2}), d_t \boldsymbol{m}_h^{i+1} \rangle \\
&\overset{(3.6)}{=} -a(\boldsymbol{m}_h^{i+1/2}, d_t \boldsymbol{m}_h^{i+1}) + \langle \boldsymbol{f}, d_t \boldsymbol{m}_h^{i+1} \rangle_\Omega \\
&\overset{(3.4)}{=} -\frac{1}{k}(\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i)).
\end{aligned}
\tag{3.37}
$$

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 (3.17). $\qquad\square$

3.5.2 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

$$a(\boldsymbol{\psi}, \boldsymbol{\varphi}) \le (C_1 + \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega); \boldsymbol{L}^2(\Omega))}) \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} \|\boldsymbol{\varphi}\|_{\boldsymbol{H}^1(\Omega)} \quad \text{for all } \boldsymbol{\psi}, \boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega), \quad (3.38\text{a})$$

$$a^{\mathrm{loc}}(\boldsymbol{\psi}, \boldsymbol{\varphi}) \le C_1 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} \|\boldsymbol{\varphi}\|_{\boldsymbol{H}^1(\Omega)} \quad \text{for all } \boldsymbol{\psi}, \boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega), \quad (3.38\text{b})$$

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

$$a(\boldsymbol{\psi}, \boldsymbol{\psi}) \ge a^{\mathrm{loc}}(\boldsymbol{\psi}, \boldsymbol{\psi}) \ge C_2 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)}^2 - C_3 \|\boldsymbol{\psi}\|_{\boldsymbol{L}^2(\Omega)}^2 \quad \text{for all } \boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega). \quad (3.38\text{c})$$

The constants C_1, C_2, C_3 in (3.38) 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 (3.15), 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 3.3.1.

Proof of Theorem 3.3.2(iii). The proof follows the lines of [BP06, PRS18b], therefore we only sketch it. Let $J \in \mathbb{N}$. Since $\boldsymbol{m}_h^J \in \boldsymbol{\mathcal{M}}_h$, $\|\boldsymbol{m}_h^J\|_{\boldsymbol{L}^2(\Omega)} \le |\Omega|^{1/2}$. Hence, combining the inequalities (3.38) and the norm equivalence (3.12) with (3.17), we obtain the estimate

$$\|\boldsymbol{m}_h^J\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \le C, \quad (3.39)$$

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 (3.39), 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 $\boldsymbol{m}_{hk}, \overline{\boldsymbol{m}}_{hk} \overset{\star}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(\mathbb{R}_{>0}; \boldsymbol{H}^1(\Omega))$, $\boldsymbol{m}_{hk}|_{\Omega_T}, \overline{\boldsymbol{m}}_{hk}|_{\Omega_T} \overset{\star}{\rightharpoonup} \boldsymbol{m}|_{\Omega_T}$ 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 [PRS18b, 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 (3.9), let $\boldsymbol{\varphi} \in \boldsymbol{C}^\infty(\overline{\Omega_T})$. Let $J \in \mathbb{N}$ 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 $\boldsymbol{\varphi}_h(t) = \boldsymbol{\mathcal{I}}_h[\boldsymbol{\varphi}(t)]$ for all $t \in [0, kJ]$. For $i = 0, \dots, J-1$ and $t \in (t_i, t_{i+1})$, we test (3.16) 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 $i = 0, \dots, J-1$, we obtain that

$$\int_0^{kJ} \langle \partial_t \boldsymbol{m}_{hk}(t), \boldsymbol{\varphi}_h(t) \rangle_h \mathrm{d}t = -\int_0^{kJ} \langle \overline{\boldsymbol{m}}_{hk}(t) \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), \boldsymbol{\varphi}_h(t) \rangle_h \mathrm{d}t$$
$$+ \alpha \int_0^{kJ} \langle \overline{\boldsymbol{m}}_{hk}(t) \times \partial_t \boldsymbol{m}_{hk}(t), \boldsymbol{\varphi}_h(t) \rangle_h \mathrm{d}t. \quad (3.40)$$

The argument in [BP06, Section 3] shows that

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

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

For the first term on the right-hand side of (3.40) simple algebraic manipulations together with (3.13) show that

$$\int_0^{kJ} \langle \overline{\boldsymbol{m}}_{hk}(t) \times \mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), \boldsymbol{\varphi}_h(t) \rangle_h \mathrm{d}t = \int_0^{kJ} \langle \boldsymbol{h}_{\text{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), (\boldsymbol{\mathcal{I}}_h - 1)[\boldsymbol{\varphi}_h(t) \times \overline{\boldsymbol{m}}_{hk}(t)] \rangle \mathrm{d}t$$
$$+ \int_0^{kJ} \langle \boldsymbol{h}_{\text{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), \boldsymbol{\varphi}_h(t) \times \overline{\boldsymbol{m}}_{hk}(t) \rangle \mathrm{d}t. \quad (3.41)$$

Since

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

and $\|(\boldsymbol{\mathcal{I}}_h - 1)[\boldsymbol{\varphi}_h \times \overline{\boldsymbol{m}}_{hk}]\|_{L^\infty(0,T;\boldsymbol{H}^1(\Omega))} \leq Ch$ (see [PRS18b, equations (39)–(40)]), the first term on the right-hand side of (3.41) 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

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

Hence, passing (3.40) to the limit as $h, k \to 0$, we obtain (3.9) for any smooth test function $\boldsymbol{\varphi}$. By density, we obtain the desired result.

Finally, the energy inequality (3.10) is obtained by passing to the limit as $h, k \to 0$ the discrete energy identity (3.17) and using standard lower semicontinuity arguments in combination with the available convergence results. $\qquad \square$

3.6 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

$$\|\boldsymbol{\nabla} \boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \leq C_{\text{inv}} h^{-1} \|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{\phi}_h \in \boldsymbol{V}_h, \quad (3.42a)$$

from which it follows that

$$\|\mathbb{P}_h \boldsymbol{\phi}\|_h \leq (1 + C_{\text{inv}}^2 h^{-2})^{1/2} \|\boldsymbol{\phi}\|_{\boldsymbol{H}^1(\Omega)^\star} \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega)^\star. \quad (3.42b)$$

Here, $C_{\mathrm{inv}} > 0$ depends only on κ. Moreover, the following inequalities are direct consequences of (3.38a)–(3.38b):

$$\|\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\psi}) - \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\varphi})\|_{\boldsymbol{H}^1(\Omega)^\star} \leq C_1 \|\boldsymbol{\psi} - \boldsymbol{\varphi}\|_{\boldsymbol{H}^1(\Omega)} \qquad \text{for all } \boldsymbol{\psi}, \boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega), \quad \text{(3.42c)}$$

$$\|\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\psi})\|_{\boldsymbol{H}^1(\Omega)^\star} \leq C_1 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \qquad \text{for all } \boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega), \quad \text{(3.42d)}$$

$$\|\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\psi}) - \boldsymbol{f}\|_{\boldsymbol{H}^1(\Omega)^\star} \leq C_1 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} \qquad \text{for all } \boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega). \quad \text{(3.42e)}$$

3.6.1 Well-posedness

We now prove Proposition 3.3.5, which establishes the properties of the constraint preserving fixed-point iteration proposed in Section 3.3.3.1.

Proof of Proposition 3.3.5(i). Since the bilinear form on the left-hand side of (3.24) 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 (3.24) 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 $\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$. $\qquad\square$

Proof of Proposition 3.3.5(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 (3.24), we obtain that

$$\begin{aligned}
&\langle \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}, \boldsymbol{\phi}_h \rangle_h \\
&\overset{(3.24)}{=} -\frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+2} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}), \boldsymbol{\phi}_h \rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}), \boldsymbol{\phi}_h \rangle_h \\
&\quad - \frac{k}{2}\langle (\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}) \times \mathbb{P}_h \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}), \boldsymbol{\phi}_h \rangle_h - \alpha\langle (\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}) \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h.
\end{aligned}$$

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

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

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

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

It follows that

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

where the last inequality follows from $\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$ and (3.42b). Moreover, it holds that

$$
\begin{aligned}
\|\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}\|_h &\leq \frac{k(1 + C_{\mathrm{inv}}^2 h^{-2})^{1/2}}{2}\|\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell})\|_{\boldsymbol{H}^1(\Omega)^\star} \\
&\overset{(3.42c)}{\leq} \frac{C_1 k(1 + C_{\mathrm{inv}}^2 h^{-2})^{1/2}}{2}\|\boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell}\|_{\boldsymbol{H}^1(\Omega)} \\
&\overset{(3.42a)}{\leq} \frac{C_1 k(1 + C_{\mathrm{inv}}^2 h^{-2})}{2}\|\boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell}\|_h.
\end{aligned}
$$

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 (3.26) holds for all $h < h_0$ and $k < k_0$. $\qquad\square$

Proof of Proposition 3.3.5(iii). Let $\ell \in \mathbb{N}_0$. Using (3.42b), (3.42c), and (3.12) as well as the fact that $\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$, we obtain that

$$
\begin{aligned}
\|\boldsymbol{\mathcal{I}}_h[\boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}))]\|_h &\leq \|\mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}))\|_h \\
&\leq C_1(1 + C_{\mathrm{inv}}^2 h^{-2})\|\boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell}\|_h \\
&\overset{(3.26)}{\leq} C_1(1 + C_{\mathrm{inv}}^2 h^{-2})q^\ell\|\boldsymbol{\eta}_h^{i,1} - \boldsymbol{\eta}_h^{i,0}\|_h \\
&\leq 2C_1|\Omega|^{1/2}(1 + C_{\mathrm{inv}}^2 h^{-2})q^\ell.
\end{aligned}
$$

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

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

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 (3.24) it follows that $\boldsymbol{m}_{h\varepsilon}^{i+1}$ solves

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

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 $\boldsymbol{m}_{h\varepsilon}^{i+1} \in \boldsymbol{\mathcal{M}}_h$ (see the proof of Theorem 3.3.2(i)). $\qquad\square$

3.6.2 Stability and weak convergence

Next, we provide the proof of Theorem 3.3.7, which establishes the stability and convergence of Algorithm 3.3.6.

Proof of Theorem 3.3.7. Part (i) is a direct consequence of Proposition 3.3.5. The proof of part (ii) follows the lines of the one of Theorem 3.3.2(ii).

Let us now consider the proof of part (iii). Testing (3.27) 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 (3.37) 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 $1 \leq j \leq J$. 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 (3.38c) and continuity (3.38a) 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

$$C_2 \|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}^2 + 2\alpha k \sum_{i=0}^{j-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2$$

$$\leq (C_1 + \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega), \boldsymbol{L}^2(\Omega))}) \|\boldsymbol{m}_h^0\|_{\boldsymbol{H}^1(\Omega)}^2 + 2\|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} (\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)})$$

$$+ C_3 \|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{L}^2(\Omega)}^2 - 2k \sum_{i=0}^{j-1} \langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h$$

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

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

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

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 (3.12), 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

$$\|\mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2})\|_h \overset{(3.42\text{b})}{\leq} (1 + C_{\text{inv}}^2 h^{-2})^{1/2} \|\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2})\|_{\boldsymbol{H}^1(\Omega)^\star}$$

$$\overset{(3.42\text{d})}{\leq} \big(1 + C_{\text{inv}}^2 h^{-2}\big)^{1/2}(C_1 \|\boldsymbol{m}_{h\varepsilon}^{i+1/2}\|_{\boldsymbol{H}^1(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}),$$

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

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

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

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 $\|\boldsymbol{\mathcal{I}}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \leq \varepsilon$ of Algorithm 3.3.6, if h is sufficiently small, we obtain that

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

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

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

where the constant $C' > 0$ depends only on T, $|\Omega|$, $\boldsymbol{f}$, κ, $\boldsymbol{\pi}$, 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 κ, 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., [Tho06, Lemma 10.5]) and the norm equivalence (3.12) yield (3.29). This concludes the proof of part (iii).

The proof of part (iv) follows the lines of [BP06, Bar06, Cim09, PRS18b]; see also the proof of Theorem 3.3.2(iii). In particular, (3.9) and (3.10) are obtained by passing to the limit as $h, k, \varepsilon \to 0$ the discrete identities (3.27) and (3.28), 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 3.3.2(iii), vanish in the limit, because they are bounded by ε (resp., because $\boldsymbol{\pi}_h$ is assumed to be consistent with $\boldsymbol{\pi}$). $\qquad\square$

3.7 Analysis of the practical midpoint scheme: Newton iteration

3.7.1 Stability of Algorithm 3.3.8

Lemma 3.3.9 and Theorem 3.3.10 assume well-posedness of Algorithm 3.3.8 up to time-step $i < J$, i.e., that for all $n = 0, \ldots, i$ the Newton solver (3.30) returns after finitely many iterations the solutions $\boldsymbol{m}_{h\varepsilon}^{n+1}, \boldsymbol{r}_{h\varepsilon}^n \in \boldsymbol{V}_h$ such that (3.32) holds with $\|\boldsymbol{r}_{h\varepsilon}^n\|_h \leq \varepsilon$. Later, in Sections 3.7.4.1–3.7.4.4 Theorem 3.3.11 is proved, guaranteeing that, given appropriate CFL-conditions, this well-posedness assumption is always satisfied.

3.7.1.1 Boundedness of magnetization length, Lemma 3.3.9(i)–(ii)

For $0 \leq n \leq i$ and $\boldsymbol{z} \in \mathcal{N}_h$, testing (3.32) with $\phi_h = \boldsymbol{m}_{h\varepsilon}^{n+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}} \in \boldsymbol{V}_h$ yields

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

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

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

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

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

$$\leq \exp(4C_{h\varepsilon}T)\left(|\boldsymbol{m}_h^0(\boldsymbol{z})|^2 + C_{h\varepsilon}T\right).$$

Using $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ 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

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

by a similar computation one derives an estimate below via

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

$$\geq \exp(-8C_{h\varepsilon}T)|\boldsymbol{m}_h^0(\boldsymbol{z})|^2 - \exp(-4C_{h\varepsilon}T)C_{h\varepsilon}T \qquad \text{for all } 0 < k < k_0,$$

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. $\qquad\square$

3.7.1.2 Stability and weak convergence, Theorem 3.3.10(i)–(iii)

For $0 \leq i < J$ testing (3.32) 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

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

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

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 3.6.2, where Theorem 3.3.7(ii)–(iv) is proved. Hence, using $\boldsymbol{L}^\infty(\Omega)$-boundedness of the iterates from Lemma 3.3.9(i) and that the stopping criterion (3.31) guarantees $\|\boldsymbol{r}_{h\varepsilon}^i\|_h \leq \varepsilon$, the proof of Theorem 3.3.10(i)–(iii) directly follows the lines of Section 3.6.2. $\qquad\square$

3.7.2 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 $\boldsymbol{F}\colon \mathbb{R}^n \to \mathbb{R}^n$,

$$\text{find } \boldsymbol{x}^* \in \mathbb{R}^n, \text{ such that } \boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{0}. \tag{3.43}$$

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 $\boldsymbol{\nabla F}(\boldsymbol{x}) \in \mathbb{R}^{n \times n}$. Given a starting value $\boldsymbol{x}^0 \in \mathbb{R}^n$, Newton's method applied to (3.43) iterates for all $\ell \in \mathbb{N}_0$

$$\begin{aligned} \text{solve} \quad & \boldsymbol{\nabla F}(\boldsymbol{x}^\ell)\boldsymbol{\delta x}^\ell = -\boldsymbol{F}(\boldsymbol{x}^\ell)\,, \\ \text{set} \quad & \boldsymbol{x}^{\ell+1} = \boldsymbol{x}^\ell + \boldsymbol{\delta x}^\ell\,. \end{aligned} \tag{3.44}$$

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 [QSS07, 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 $\|\boldsymbol{A}\boldsymbol{x}\|_{\mathbb{R}^n} \le \|\boldsymbol{A}\|_{\mathbb{R}^{n\times n}} \|\boldsymbol{x}\|_{\mathbb{R}^n}$ 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

$$\|\boldsymbol{A}\| = \sup_{\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \quad \text{for all } \boldsymbol{A} \in \mathbb{R}^{n\times n}. \tag{3.45}$$

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

Theorem 3.7.1 ([QSS07, 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 $(\boldsymbol{\nabla}\boldsymbol{F}(\boldsymbol{x}^*))^{-1} \in \mathbb{R}^{n\times n}$ exists and that there exist constants $C, R, L > 0$, such that*

$$\|(\boldsymbol{\nabla}\boldsymbol{F}(\boldsymbol{x}^*))^{-1}\| \le C, \tag{3.46a}$$

and

$$\|\boldsymbol{\nabla}\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{\nabla}\boldsymbol{F}(\boldsymbol{y})\| \le L\|\boldsymbol{x} - \boldsymbol{y}\| \quad \textit{for all} \quad \boldsymbol{x}, \boldsymbol{y} \in B(\|\cdot\|; \boldsymbol{x}^*, R), \tag{3.46b}$$

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 (3.44) is uniquely defined and converges to $\boldsymbol{x}^*$ with

$$\|\boldsymbol{x}^{\ell+1} - \boldsymbol{x}^*\| \le CL\|\boldsymbol{x}^\ell - \boldsymbol{x}^*\|^2. \tag{3.47}$$

Remark 3.7.2. *In particular* (3.47) *and $\boldsymbol{x}^0 \in B(\|\cdot\|; \boldsymbol{x}^*, \min\{R, 1/(2CL)\})$ imply*

$$\|\boldsymbol{x}^\ell - \boldsymbol{x}^*\| \le \left(\prod_{j=0}^{\ell-1} (CL)^{2^j}\right) \|\boldsymbol{x}^0 - \boldsymbol{x}^*\|^{2^\ell} = (CL)^{2^\ell-1}\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|^{2^\ell} \le \frac{2\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|}{2^{2^\ell}}, \tag{3.48}$$

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

3.7.3 Newton's method applied to the nonlinear midpoint scheme

We aim to apply Newton's method (3.44) to the nonlinear system of equations (3.23), i.e., to the IMEX version of the ideal midpoint scheme where the lower order terms are integrated explicitly in time $\boldsymbol{\pi}(\boldsymbol{m}_h^{i+1/2}) \approx \boldsymbol{\Pi}_h(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1})$. Consider a numbering of the nodes $\{\boldsymbol{z}_j \colon j = 1, \ldots, N\} = \mathcal{N}_h$ of the mesh $\mathcal{T}_h$, and associate with a given vector $\boldsymbol{x} \in (\mathbb{R}^3)^N$ the finite element function defined by $\widehat{\boldsymbol{x}} := \sum_{j=1}^N \boldsymbol{x}_j \varphi_{\boldsymbol{z}_j} \in \boldsymbol{V}_h$. Further, for a finite element function $\boldsymbol{u} = \sum_{j=1}^N \boldsymbol{u}(\boldsymbol{z}_j)\varphi_{\boldsymbol{z}_j} \in \boldsymbol{V}_h$, we write $[\boldsymbol{u}] \in (\mathbb{R}^3)^N \simeq \mathbb{R}^{3N}$ for the vector of nodal values, i.e., $[\boldsymbol{u}]_j := \boldsymbol{u}(\boldsymbol{z}_j) \in \mathbb{R}^3$.

The mass lumped scalar product $\langle \cdot, \cdot \rangle_h$ gives rise to the matrix $\mathbf{M}_h \in (\mathbb{R}^{3\times 3})^{N\times N} \simeq \mathbb{R}^{3N\times 3N}$, defined via $(\mathbf{M}_h)_{jk} := \delta_{j,k}\beta_{\boldsymbol{z}_j}\boldsymbol{I}_{3\times 3} \in \mathbb{R}^{3\times 3}$. Given $\boldsymbol{m}_h^i \in \boldsymbol{V}_h$, the solution $\boldsymbol{m}_h^{i+1/2}$ of (3.23) satisfies $\boldsymbol{F}([\boldsymbol{m}_h^{i+1/2}]) = \boldsymbol{0}$, with

$$\boldsymbol{F}(\boldsymbol{x}) := \mathbf{M}_h\left(\boldsymbol{x} - [\boldsymbol{m}_h^i] + \left[\boldsymbol{\mathcal{I}}_h\left(\frac{k}{2}\widehat{\boldsymbol{x}} \times \mathbb{P}_h\left(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\widehat{\boldsymbol{x}}) + \boldsymbol{\Pi}_h(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1})\right) + \alpha\widehat{\boldsymbol{x}} \times \boldsymbol{m}_h^i\right)\right]\right). \tag{3.49}$$

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

$$
\boldsymbol{u}^\top \boldsymbol{\nabla F}(\boldsymbol{x}) \boldsymbol{v} = \langle \widehat{\boldsymbol{u}}, \widehat{\boldsymbol{v}} \rangle_h + \frac{k}{2} \langle \widehat{\boldsymbol{u}} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{x}}), \widehat{\boldsymbol{v}} \rangle_h + \frac{k}{2} \langle \widehat{\boldsymbol{x}} \times \mathbb{P}_h \big(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{u}}) - \boldsymbol{f} \big), \widehat{\boldsymbol{v}} \rangle_h
$$
$$
+ \frac{k}{2} \langle \widehat{\boldsymbol{u}} \times \mathbb{P}_h \boldsymbol{\Pi}_h(\boldsymbol{m}_h^i, \boldsymbol{m}_h^{i-1}), \widehat{\boldsymbol{v}} \rangle_h + \alpha \langle \widehat{\boldsymbol{u}} \times \boldsymbol{m}_h^i, \widehat{\boldsymbol{v}} \rangle_h \,. \tag{3.50}
$$

Newton's method (3.44) applied to the system (3.23) in the ith time-step now can be written as: Given $\boldsymbol{m}_{h\varepsilon}^i \in \boldsymbol{V}_h$ and initial value $\boldsymbol{\eta}_h^{i,0} \in \boldsymbol{V}_h$, for all $\ell \in \mathbb{N}_0$ compute $\boldsymbol{u}_h^{i,\ell} \in \boldsymbol{V}_h$ such that, for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$, it holds that

$$
\langle \boldsymbol{u}_h^{i,\ell}, \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \boldsymbol{u}_h^{i,\ell} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}), \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell} \times \mathbb{P}_h \big(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{u}_h^{i,\ell}) - \boldsymbol{f} \big), \boldsymbol{\phi}_h \rangle_h
$$
$$
+ \frac{k}{2} \langle \boldsymbol{u}_h^{i,\ell} \times \mathbb{P}_h \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \boldsymbol{u}_h^{i,\ell} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h
$$
$$
= \langle \boldsymbol{m}_{h\varepsilon}^i - \boldsymbol{\eta}_h^{i,\ell}, \boldsymbol{\phi}_h \rangle_h - \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}), \boldsymbol{\phi}_h \rangle_h \tag{3.51a}
$$
$$
- \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell} \times \mathbb{P}_h \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}), \boldsymbol{\phi}_h \rangle_h - \alpha \langle \boldsymbol{\eta}_h^{i,\ell} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h \,,
$$

and define

$$
\boldsymbol{\eta}_h^{i,\ell+1} := \boldsymbol{\eta}_h^{i,\ell} + \boldsymbol{u}_h^{i,\ell} \,. \tag{3.51b}
$$

In the remainder of this section, to improve readability we omit the h-subscript of the iteration variables $\boldsymbol{\eta}^{i,\ell}$ and $\boldsymbol{u}^{i,\ell}$. Note that by (3.49)–(3.50) we see that (3.51a)–(3.51b) resembles Newton's method (3.44) with $\boldsymbol{x}^\ell = [\boldsymbol{\eta}^{i,\ell}]$ and $\boldsymbol{\delta x}^\ell = [\boldsymbol{u}^{i,\ell}]$. Given some tolerance $\varepsilon > 0$, the iteration will be stopped once

$$
\| \mathcal{I}_h \big(\boldsymbol{u}^{i,\ell} \times \mathbb{P}_h \big(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{u}^{i,\ell}) - \boldsymbol{f} \big) \big) \|_h \leq \varepsilon. \tag{3.52}
$$

If $\ell^* \in \mathbb{N}_0$ is the first index for which the stopping criterion (3.52) is satisfied, the approximate magnetization at the new time-step is defined as $\boldsymbol{m}_{h\varepsilon}^{i+1} := 2\,\boldsymbol{\eta}^{i,\ell^*+1} - \boldsymbol{m}_{h\varepsilon}^i$.

For all $i \in \mathbb{N}_0$, let $\boldsymbol{r}_{h\varepsilon}^i := \mathcal{I}_h \big(\boldsymbol{u}^{i,\ell^*} \times \mathbb{P}_h \big(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{u}^{i,\ell^*}) - \boldsymbol{f} \big) \big) \in \boldsymbol{V}_h$, so that $\langle \boldsymbol{r}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h$ equals the difference of (3.23) and (3.51a). In view of the stopping criterion (3.52), it holds that $\| \boldsymbol{r}_{h\varepsilon}^i \|_h \leq \varepsilon$. With this definition, the proposed linearization of one iteration of Algorithm 3.3.1 based on the Newton method is covered by Algorithm 3.3.8.

3.7.4 Well-posedness of Algorithm 3.3.8

We show Theorem 3.3.11(i) by induction: For $0 \leq i < J$ assume that Algorithm 3.3.8 is well-defined for all $n = 0, \ldots, i-1$. In particular, by Lemma 3.3.9 and Theorem 3.3.10 we have the bounds

$$
\| \boldsymbol{m}_{h\varepsilon}^i \|_{\boldsymbol{L}^\infty(\Omega)} \leq C_\infty \quad \text{and} \quad \mathcal{E}(\boldsymbol{m}_{h\varepsilon}^i) \leq C \,. \tag{3.53}
$$

Now, the inductive step is to prove convergence of the Newton iteration (3.30)–(3.31) for time-step $n = i$. We do this by verifying the assumptions (3.46a)–(3.46b) of Theorem 3.7.1 for the Newton solver (3.51a) with the initial value chosen as $\boldsymbol{\eta}^{i,0} := \boldsymbol{m}_{h\varepsilon}^i$. In Section 3.7.4.1 we verify the Lipschitz continuity (3.46b). Invertibility (3.46a) is shown in

Section 3.7.4.2. In Section 3.7.4.3 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 3.7.4.4 we conclude by estimating the maximum number of Newton iterations required to achieve the required tolerance (3.52), in particular showing that the number is finite. Hence, Sections 3.7.4.1–3.7.4.4 prove Theorem 3.3.11.

Throughout the proof, we use the notation of Section 3.7.2–3.7.3 and consider the ℓ^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. (3.45).

3.7.4.1 Lipschitz continuity of $\boldsymbol{\nabla F}$

By (3.45) it holds for arbitrary $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{3N}$ that

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

With the representation (3.50) and the estimates (3.42) we see

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

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

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

In particular, (3.46b) holds for $\|\cdot\|_2$ with $R = +\infty$ and $L \simeq kh$. $\qquad\square$

3.7.4.2 Invertibility of $\boldsymbol{\nabla F}(\boldsymbol{x}^*)$

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

$$0 = [\widehat{\boldsymbol{x}^*}(\boldsymbol{z}_j)\varphi_{\boldsymbol{z}_j}]^\top \boldsymbol{F}(\boldsymbol{x}^*) \overset{(3.49)}{=} \beta_{\boldsymbol{z}_j}\big(|\boldsymbol{x}_j^*|^2 - \boldsymbol{x}_j^* \cdot \boldsymbol{m}_{h\varepsilon}^i(\boldsymbol{z}_j)\big) \quad \text{for all} \quad j = 1, \dots, N$$

together with (3.53) guarantees boundedness

$$\|\widehat{\boldsymbol{x}^*}\|_{\boldsymbol{L}^\infty(\Omega)} \leq \|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{L}^\infty(\Omega)} \leq C_\infty\,. \tag{3.54}$$

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

$$\boldsymbol{y}^\top \boldsymbol{\nabla F}(\boldsymbol{x}^*)\boldsymbol{y} \overset{(3.50)}{=} \|\widehat{\boldsymbol{y}}\|_h^2 + \frac{k}{2}\langle \widehat{\boldsymbol{x}^*} \times \mathbb{P}_h\big(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\widehat{\boldsymbol{y}}) - \boldsymbol{f}\big), \widehat{\boldsymbol{y}}\rangle_h \gtrsim (1 - kh^{-2})\|\widehat{\boldsymbol{y}}\|_h^2 \gtrsim \|\widehat{\boldsymbol{y}}\|_h^2\,,$$

where we used (3.54) and the Cauchy–Schwarz inequality together with the estimates (3.42). To show boundedness of $(\boldsymbol{\nabla}\boldsymbol{F}(\boldsymbol{x}^*))^{-1}$ we write

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

with $\boldsymbol{y} := \boldsymbol{y}(\boldsymbol{x}) := (\boldsymbol{\nabla}\boldsymbol{F}(\boldsymbol{x}^*))^{-1}\boldsymbol{x}$. Using (3.50), it holds that

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

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

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

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

$$\overset{(3.42)}{\lesssim} \|\boldsymbol{y}\|_2\|\boldsymbol{x}\|_2 + kh^{-2}\|\widehat{\boldsymbol{y}}\|_h^2 \lesssim \|\boldsymbol{y}\|_2\|\boldsymbol{x}\|_2 + kh\|\boldsymbol{y}\|_2^2.$$

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

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

In particular it holds (3.46a) for $\|\cdot\|_2$ with $C \simeq h^{-3}$. $\qquad\qquad\square$

3.7.4.3 Initial guess leads to convergence

We recall the results from Section 3.7.4.1 and Section 3.7.4.2: The Newton iteration (3.51a) satisfies the assumptions of Theorem 3.7.1 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 (3.51a) for any initial guess $\boldsymbol{x}^0 \in \mathbb{R}^{3N}$ 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 3.3.8 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 (3.43), i.e., by (3.49) it holds for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$ that

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

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

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

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

We rewrite the first term on the right hand side as

$$
4\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\widehat{\boldsymbol{x}^*}) \rangle_h \overset{(3.13)}{=} 4\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{h}_{\mathrm{eff}}(\widehat{\boldsymbol{x}^*}) \rangle
$$
$$
\overset{(3.6)}{=} -4a(\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \widehat{\boldsymbol{x}^*}) + 4\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f} \rangle_\Omega
$$
$$
= -a\big((2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) - \boldsymbol{m}_{h\varepsilon}^i, (2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) + \boldsymbol{m}_{h\varepsilon}^i \big)
$$
$$
\quad + 2\langle 2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f} \rangle_\Omega - 2\langle \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f} \rangle_\Omega
$$
$$
\overset{(3.4)}{=} 2\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^i) - a(2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, 2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) + 2\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f} \rangle_\Omega .
$$

With (3.54) and the Gårding inequality (3.38c) we estimate

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

Now combination with the generous estimates

$$
\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \mathbb{P}_h \big(\boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}) - \boldsymbol{\pi}(\widehat{\boldsymbol{x}^*}) \big) \rangle_h \le 2C_\infty^2 |\Omega| (2C_{\boldsymbol{\pi}} + \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega);\boldsymbol{L}^2(\Omega))}) ,
$$
$$
\langle 2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f} \rangle_\Omega \le 3C_\infty \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}
$$

yields

$$
\|\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i\|_h^2 \lesssim \mathcal{E}(\boldsymbol{m}_{h\varepsilon}^i) k + C\big(C_3, C_\infty, |\Omega|, C_{\boldsymbol{\pi}}, \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega);\boldsymbol{L}^2(\Omega))}, \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \big) k \overset{(3.53)}{\lesssim} k .
$$

Due to the norm equivalence $h^{3/2} \|\cdot\|_2 \simeq \|\widehat{\cdot}\|_h$ 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 $\ell \to \infty$. $\qquad\square$

3.7.4.4 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 (3.52) is satisfied. The index ℓ^* is well defined due to $\widehat{\boldsymbol{\delta x}^\ell} = \boldsymbol{u}^{i,\ell}$ and

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

Recalling that by Remark 3.7.2 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 the estimates (3.42) and the norm equivalence $h^{3/2} \|\cdot\|_2 \simeq \|\widehat{\cdot}\|_h$ on $\mathbb{R}^{3N}$, it holds for the error $\boldsymbol{r}_{h\varepsilon}^i$ that

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

Since $\ell^* \in \mathbb{N}$ is defined as the smallest integer, such that (3.52) is satisfied, ℓ^* is estimated from above by $\log_2 \log_2(C_\star k h^{-7/2} \varepsilon^{-1})$ with a generic constant $C_\star > 0$. □

4 Iterative solution and preconditioning for the tangent plane scheme in computational micromagnetics

This chapter consists of the article [KPP$^+$19] together with Johannes Kraus, Dirk Praetorius, Michele Ruggeri, and Bernhard Stiftner.

Abstract. The tangent plane scheme is a time-marching scheme for the numerical solution of the nonlinear parabolic Landau–Lifshitz–Gilbert equation, which describes the time evolution of ferromagnetic configurations. Exploiting the geometric structure of the equation, the tangent plane scheme requires only the solution of one linear variational form per time-step, which is posed in the discrete tangent space determined by the nodal values of the current magnetization. We develop an effective solution strategy for the arising constrained linear systems, which is based on appropriate Householder reflections. We derive possible preconditioners, which are (essentially) independent of the time-step, and prove linear convergence of the preconditioned GMRES algorithm. Numerical experiments underpin the theoretical findings.

4.1 Introduction

4.1.1 Landau–Lifshitz–Gilbert equation

The Landau–Lifshitz–Gilbert equation (LLG) describes time-dependent micromagnetic phenomena in a bounded ferromagnetic domain $\Omega \subset \mathbb{R}^3$ with Lipschitz boundary $\partial\Omega$ [Gil55, LL08]. After a suitable scaling of the involved physical quantities, it reads

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) + \alpha\, \boldsymbol{m} \times \partial_t \boldsymbol{m} \qquad \text{in } (0,T) \times \Omega, \qquad (4.1\text{a})$$

$$\partial_{\boldsymbol{n}} \boldsymbol{m} = \boldsymbol{0} \qquad \text{on } (0,T) \times \partial\Omega, \qquad (4.1\text{b})$$

$$\boldsymbol{m}(0) = \boldsymbol{m}^0 \qquad \text{in } \Omega, \qquad (4.1\text{c})$$

where the unknown $\boldsymbol{m} : (0,T) \times \Omega \to \mathbb{R}^3$ is the magnetization, $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m})$ is the effective field, $\alpha \in (0,1]$ is the Gilbert damping constant, $T > 0$ is the final time, and $\boldsymbol{m}^0 : \Omega \to \mathbb{R}^3$ with $|\boldsymbol{m}^0| = 1$ in Ω is the initial configuration. The effective field comprises several contributions, which correspond to different phenomena in micromagnetism. In usual applications (see, e.g., [HS98]), it takes the form $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) := \ell_{\text{ex}}^2 \Delta \boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$, where $\ell_{\text{ex}}^2 \Delta \boldsymbol{m}$ is the exchange field with the exchange length $\ell_{\text{ex}} > 0$, $\boldsymbol{\pi}(\boldsymbol{m})$ is a short-hand notation for all $\boldsymbol{m}$-dependent lower-order terms (e.g., stray field or magnetocrystalline anisotropy), and $\boldsymbol{f} : \Omega \to \mathbb{R}^3$ is an applied external field.

Taking the scalar product with $\boldsymbol{m}$ in (4.1a), we note the PDE inherent constraints

$$\frac{1}{2}\,\partial_t|\boldsymbol{m}|^2 = \boldsymbol{m}\cdot\partial_t\boldsymbol{m} = 0 \quad\text{and thus}\quad |\boldsymbol{m}| = 1 \quad\text{in }(0,T)\times\Omega. \tag{4.2}$$

In particular, $\partial_t\boldsymbol{m}(t)$ belongs to the tangent space of $\boldsymbol{m}(t)$ for all $t\in(0,T)$.

4.1.2 Tangent plane schemes

Using (4.2), one can reformulate (4.1a) as

$$\alpha\,\partial_t\boldsymbol{m} + \boldsymbol{m}\times\partial_t\boldsymbol{m} = \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - (\boldsymbol{h}_{\text{eff}}(\boldsymbol{m})\cdot\boldsymbol{m})\,\boldsymbol{m}, \tag{4.3}$$

which is a linear equation in $\boldsymbol{v} := \partial_t\boldsymbol{m}$. The idea of the tangent plane scheme (TPS) can be summarized as follows: At time t_n, the magnetization $\boldsymbol{m}(t_n)$ is approximated by $\boldsymbol{m}^n$ in the lowest-order conforming finite element space. Discretizing (4.3) by a Galerkin approach in the discrete tangent space at $\boldsymbol{m}^n$, we obtain an approximation $\boldsymbol{v}^n \approx \boldsymbol{v}(t_n)$. Up to nodal normalization, $\boldsymbol{m}^n + k\boldsymbol{v}^n$ then yields an approximation of the magnetization at time $t_{n+1} := t_n + k$, where $k > 0$ is the time-step size. Although LLG is nonlinear, TPS thus requires only the solution of one linear system per time-step for $\boldsymbol{v}^n$, however, in the discrete tangent space.

TPS with explicit time-stepping was first analyzed in [AJ06] with a refined analysis in [BKP08], which requires a CFL condition for convergence towards a weak solution in the sense of [AS92]. The work [Alo08] proposed TPS with an implicit time-stepping. This yields unconditional convergence of the algorithm towards a weak solution. While [Alo08] considered only the exchange field, it was extended to general stationary lower-order contributions in [AKT12, BFF$^+$14] and chiral magnetic skyrmion dynamics in [HPP$^+$19]. Moreover, TPS was extended to the coupling of LLG with other evolution equations such as the full Maxwell system [BPP15], the eddy current equation [LT13, LPPT15, FT17a], the conservation of elastic momentum [BPPR14] when modeling magnetostrictive effects, or a spin diffusion equation [AHP$^+$14, ARB$^+$15]. In the mentioned works, TPS is formally of first order in time. Recently, TPS was modified into a (formally) second-order in time scheme in [AKST14] with extensions in [DPP$^+$20, HPP$^+$19].

4.1.3 Contributions

So far, the efficient solution of the constrained linear system in the discrete tangent space for the computation of $\boldsymbol{v}^n \approx \boldsymbol{v}(t_n)$ has not been discussed in the literature. For related works, which empirically investigate the solution of the algebraic systems arising from alternative integrators for LLG, we refer to [Baň10, STS$^+$02]. For tangent plane schemes, the main difficulty is the time-dependent ansatz space. This aggravates the construction of suitable and effective preconditioners, which, if possible, should not depend on the time-step, or, at least, only need an update every once in a while (after several time-steps).

We construct a linear system in $\mathbb{R}^{2N}$, where $N\in\mathbb{N}$ is the number of nodes of the underlying finite element discretization. The corresponding system matrix is positive definite, but non-symmetric and depends on the time-step. We present and analyze various preconditioners, including a stationary approach (i.e., independent of the time-step) as well as

Jacobi-type approximations. In the worst case, the number of necessary updates of the pre-conditioner to attain optimal convergence of the GMRES algorithm [SS86, Saa03] depends on the mesh-size h. However, under certain assumptions on the discrete magnetization $\boldsymbol{m}^n \approx \boldsymbol{m}(t_n)$, the number of necessary updates is also independent of h.

4.1.4 Outline

This paper is organized as follows: Section 4.2 introduces the basic notation and gives a precise formulation of TPS (Algorithm 4.2.1). In Section 4.3, we provide a basis for the discrete tangent space and derive the prototype linear system, which has to be solved in each time-step (Theorem 4.3.1). Section 4.4 proposes symmetric and positive definite preconditioners for the latter linear system. The two main results (Theorem 4.4.1 and Theorem 4.4.3) prove that the corresponding preconditioned GMRES algorithms converge linearly. These theorems also provide estimates of the residual reduction factors (in certain energy norms) and show under which assumptions these estimates are independent of the discretization parameters. Finally, we also discuss Jacobi-type approximations of our preconditioners (Section 4.4.4). Our theoretical results are underpinned by numerical experiments in Section 4.5. In Section 4.7 the theoretical and numerical results are summarized and concluding remarks on choosing the appropriate preconditioner are given. The proofs of Theorem 4.3.1, Theorem 4.4.1, and Theorem 4.4.3 are postponed to Section 4.6.

4.2 Preliminaries

4.2.1 General notation

For any dimension $d \in \mathbb{N}$ (clear from the context) and vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, let $\mathbf{x} \cdot \mathbf{y}$ denote the Euclidean scalar product with the corresponding norm $|\mathbf{x}|^2 := \mathbf{x} \cdot \mathbf{x}$. Moreover, we denote by $\mathbf{e}_i$ the i-th unit vector and by $\mathbf{I}$ the identity matrix in $\mathbb{R}^d$. To abbreviate notation, we follow the `Matlab` syntax: For vectors, $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^d$, we write $[\mathbf{x}_1, \ldots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ for the matrix whose j-th column is $\mathbf{x}_j$.

We use bold letters for vector-valued spaces, e.g., $\boldsymbol{L}^2(\Omega) = (L^2(\Omega))^3$. By slight abuse of notation, we write $\|\cdot\|_{\boldsymbol{L}^2(\Omega)}$ simultaneously for the L^2-norm on $(L^2(\Omega))^3$ and $(L^2(\Omega))^{3 \times 3}$.

4.2.2 Discretization

For the temporal discretization of LLG, let $M \in \mathbb{N}$ and $k := T/M$. Let $t_n := kn$ with $n \in \{0, \ldots, M\}$ be the uniform time-steps. For the spatial discretization, let $\mathcal{T}_h$ be a $C_{\mathbf{mesh}}$-quasi-uniform triangulation of Ω into tetrahedra $K \in \mathcal{T}_h$ with mesh-size $h > 0$, i.e., there exists $C_{\mathbf{mesh}} > 0$ such that

$$C_{\mathbf{mesh}}^{-1} h \leq |K|^{1/3} \leq \operatorname{diam}(K) \leq h \quad \text{for all } K \in \mathcal{T}_h,$$

where $|K|$ denotes the volume of the element K and $\operatorname{diam}(K)$ its diameter. Let

$$\boldsymbol{S}_h := (\mathcal{S}_h)^3 \quad \text{with} \quad \mathcal{S}_h := \big\{ v : \Omega \to \mathbb{R} \text{ continuous} : \text{for all } T \in \mathcal{T}_h \quad v_h|_T \text{ is affine} \big\}$$

be the lowest-order FEM space. We denote by $\mathcal{N}_h$ the set of nodes of $\mathcal{T}_h$ and define $N := \#\mathcal{N}_h$. Moreover, let

$$\mathcal{M}_h := \big\{ \boldsymbol{\varphi} \in \boldsymbol{S}_h \ : \ |\boldsymbol{\varphi}(z)| = 1 \quad \text{for all } z \in \mathcal{N}_h \big\}.$$

For some fixed $\boldsymbol{\mu} \in \mathcal{M}_h$, define the discrete tangent space

$$\mathcal{K}_h[\boldsymbol{\mu}] := \big\{ \boldsymbol{\varphi} \in \boldsymbol{S}_h \ : \ \boldsymbol{\varphi}(z) \cdot \boldsymbol{\mu}(z) = 0 \quad \text{for all } z \in \mathcal{N}_h \big\}. \tag{4.4}$$

Note that $\dim \boldsymbol{S}_h = 3N$ and $\dim \mathcal{K}_h[\boldsymbol{\mu}] = 2N$.

4.2.3 Tangent plane scheme

For ease of presentation, we focus on the implicit first-order TPS [Alo08, AKT12, BFF$^+$14]. We note that all results of our work can also be adapted to the second-order TPS [AKST14, DPP$^+$20]; see Remark 4.2.3.

Algorithm 4.2.1 (First-order TPS). *Input: $\boldsymbol{m}^0 \in \mathcal{M}_h$.*
Loop: *For all time-steps $n = 0, \ldots, M-1$, iterate the following steps (a)–(b):*
(a) Find $\boldsymbol{v}^n \in \mathcal{K}_h[\boldsymbol{m}^n]$ such that

$$
\begin{aligned}
\alpha \int_\Omega \boldsymbol{v}^n \cdot \boldsymbol{\varphi} \, dx &+ \int_\Omega (\boldsymbol{m}^n \times \boldsymbol{v}^n) \cdot \boldsymbol{\varphi} \, dx + \ell_{\text{ex}}^2 k \int_\Omega \nabla \boldsymbol{v}^n \cdot \nabla \boldsymbol{\varphi} \, dx \\
&= -\ell_{\text{ex}}^2 \int_\Omega \nabla \boldsymbol{m}^n \cdot \nabla \boldsymbol{\varphi} \, dx + \int_\Omega \big(\boldsymbol{\pi}(\boldsymbol{m}^n) + \boldsymbol{f}(t_n) \big) \cdot \boldsymbol{\varphi} \, dx \quad \text{for all } \boldsymbol{\varphi} \in \mathcal{K}_h[\boldsymbol{m}^n];
\end{aligned}
\tag{4.5}
$$

(b) Define $\boldsymbol{m}^{n+1} \in \mathcal{M}_h$ by nodal projection

$$\boldsymbol{m}^{n+1}(z) := \frac{\boldsymbol{m}^n(z) + k\boldsymbol{v}^n(z)}{|\boldsymbol{m}^n(z) + k\boldsymbol{v}^n(z)|} \quad \text{for all nodes } z \in \mathcal{N}_h. \tag{4.6}$$

Output: *Approximations $\boldsymbol{m}^n \approx \boldsymbol{m}(t_n)$ for all $n = 1, \ldots, M$.* $\qquad\qquad\square$

Remark 4.2.2. *(i) The variational formulation (4.5) of TPS relies on the equivalent formulation (4.3) of LLG (4.1), where the nonlinear term vanishes due to the choice of the test space (i.e., the tangent space).*

(ii) The bilinear form on the left-hand side of (4.5) is continuous and elliptic. Therefore, the Lax–Milgram theorem guarantees existence and uniqueness of the solution $\boldsymbol{v}^n \in \mathcal{K}_h[\boldsymbol{m}^n]$ to (4.5).

(iii) To see that (4.6) is well-defined, note that $\boldsymbol{m}^0 \in \mathcal{M}_h$ and induction on n prove that

$$|\boldsymbol{m}^n(z) + k\boldsymbol{v}^n(z)|^2 \overset{(4.4)}{=} |\boldsymbol{m}^n(z)|^2 + k^2 |\boldsymbol{v}^n(z)|^2 \geq 1 \quad \text{for all nodes } z \in \mathcal{N}_h.$$

(iv) A variant of Algorithm 4.2.1 is the projection-free TPS [AHP$^+$14], where the nodal projection is omitted, i.e., $\boldsymbol{m}^{n+1} := \boldsymbol{m}^n + k\boldsymbol{v}^n \in \boldsymbol{S}_h$ in step (b). This projection-free integrator remains unconditionally convergent [AHP$^+$14]. Moreover, if there exists a smooth (and hence unique [DS14]) strong solution of LLG, the a priori analysis in [FT17a] guarantees first-order convergence in space and time for this projection-free integrator.

Remark 4.2.3. *For the second-order TPS [AKST14, DPP⁺20], the left-hand side of* (4.5) *takes the form*

$$\int_\Omega W_k(\lambda^n) \boldsymbol{v}^n \cdot \boldsymbol{\varphi} \, dx + \int_\Omega (\boldsymbol{m}^n \times \boldsymbol{v}^n) \cdot \boldsymbol{\varphi} \, dx + \ell_{\mathrm{ex}}^2 \frac{1 + \rho(k)}{2} k \int_\Omega \nabla \boldsymbol{v}^n \cdot \nabla \boldsymbol{\varphi} \, dx.$$

Here, $\lambda^n := -\ell_{\mathrm{ex}}^2 |\nabla \boldsymbol{m}^n|^2 + \big(\boldsymbol{\pi}(\boldsymbol{m}^n) + \boldsymbol{f}(t_n)\big) \cdot \boldsymbol{m}^n$, $\rho(k) := |k \log(k)|$, *and*

$$W_k(s) := \begin{cases} \alpha + (k/2) \min\{s, 1/\rho(k)\} & \text{for } s \geq 0, \\ 2\alpha^2/[2\alpha + k \min\{-s, 1/\rho(k)\}] & \text{for } s < 0. \end{cases}$$

We note that $W_k(\cdot) \to \alpha$ *as* $k \to 0$. *Moreover, the right-hand side of* (4.5) *requires minor modifications to be second-order accurate; see [AKST14, DPP⁺20] for details. We refer to the extended preprint [KPP⁺18] of this work, where the extension of our results to the second-order TPS is discussed in detail.*

4.2.4 Linear algebra

We suppose a numbering of the nodes, i.e., $\mathcal{N}_h = \{z_1, \ldots, z_N\}$. Let $\varphi_j \in \mathcal{S}_h$ be the nodal hat function associated with z_j, i.e., $\varphi_j(z_j) = 1$ and $\varphi_j(z_i) = 0$ for $i \neq j$. We then consider the following basis of $\boldsymbol{\mathcal{S}}_h$: Define

$$\boldsymbol{\phi}_{3(j-1)+\ell} := \varphi_j \, \boldsymbol{e}_\ell : \Omega \to \mathbb{R}^3 \quad \text{for all } j = 1, \ldots, N \text{ and all } \ell = 1, 2, 3. \tag{4.7}$$

Given $\boldsymbol{m} \in \boldsymbol{\mathcal{M}}_h$, we then define $\mathbf{M}, \mathbf{L}, \mathbf{S}[\boldsymbol{m}] \in \mathbb{R}^{3N \times 3N}$ as follows:

- $\mathbf{M}_{ij} := \int_\Omega \boldsymbol{\phi}_j \cdot \boldsymbol{\phi}_i \, dx$ is the (symmetric, positive definite) mass matrix;

- $\mathbf{L}_{ij} := \int_\Omega \nabla \boldsymbol{\phi}_j \cdot \nabla \boldsymbol{\phi}_i \, dx$ is the (symmetric, positive semidefinite) stiffness matrix;

- $(\mathbf{S}[\boldsymbol{m}])_{ij} := \int_\Omega (\boldsymbol{m} \times \boldsymbol{\phi}_j) \cdot \boldsymbol{\phi}_i \, dx$ is the (skew-symmetric) cross product matrix.

Moreover, we set

$$M_{ij} := \int_\Omega \varphi_j \varphi_i \, dx \in \mathbb{R} \quad \text{and} \quad L_{ij} := \int_\Omega \nabla \varphi_j \cdot \nabla \varphi_i \, dx \in \mathbb{R} \quad \text{for } i, j = 1, \ldots, N, \tag{4.8a}$$

and note the block forms

$$\mathbf{M} = \begin{pmatrix} M_{11} \mathbf{I}_{3\times3} & \cdots & M_{1N} \mathbf{I}_{3\times3} \\ \vdots & \ddots & \vdots \\ M_{N1} \mathbf{I}_{3\times3} & \cdots & M_{NN} \mathbf{I}_{3\times3} \end{pmatrix} \quad \text{and} \quad \mathbf{L} = \begin{pmatrix} L_{11} \mathbf{I}_{3\times3} & \cdots & L_{1N} \mathbf{I}_{3\times3} \\ \vdots & \ddots & \vdots \\ L_{N1} \mathbf{I}_{3\times3} & \cdots & L_{NN} \mathbf{I}_{3\times3} \end{pmatrix}. \tag{4.8b}$$

If we replace $\mathcal{K}_h[\boldsymbol{m}^n]$ with $\boldsymbol{\mathcal{S}}_h$ in (4.5), the left-hand side of (4.5) gives rise to the matrix

$$\mathbf{A}[\boldsymbol{m}^n] := \alpha \mathbf{M} + \mathbf{S}[\boldsymbol{m}^n] + \ell_{\mathrm{ex}}^2 k \mathbf{L} \in \mathbb{R}^{3N \times 3N}. \tag{4.9}$$

The right-hand side of (4.5) gives rise to the vector $\mathbf{b}[\boldsymbol{m}^n] \in \mathbb{R}^{3N}$ with

$$\big(\mathbf{b}[\boldsymbol{m}^n]\big)_j := -\ell_{\mathrm{ex}}^2 \int_\Omega \nabla \boldsymbol{m}^n \cdot \nabla \boldsymbol{\phi}_j \, dx + \int_\Omega \big(\boldsymbol{\pi}(\boldsymbol{m}^n) + \boldsymbol{f}(t_n)\big) \cdot \boldsymbol{\phi}_j \, dx.$$

Note that the matrix $\mathbf{A}[\boldsymbol{m}^n]$ is positive definite and hence regular, but not symmetric.

4.3 The tangent space problem

In this section, we present a strategy, which translates the solution of the discrete variational formulation (4.5) into a linear system in $\mathbb{R}^{2N} \cong \mathcal{K}_h[\boldsymbol{m}^n]$. To that end, we use Householder matrices: Given $\mathbf{w} = \boldsymbol{m}^n(z) \in \mathbb{R}^3$ with $|\mathbf{w}| = 1$, define $\widetilde{\mathbf{H}}[\mathbf{w}] \in \mathbb{R}^{3\times3}$ by

$$\widetilde{\mathbf{H}}[\mathbf{w}] := \begin{cases} \mathbf{I} - 2\mathbf{z}\mathbf{z}^T, & \text{where} \quad \mathbf{z} := \frac{\mathbf{w}+\mathbf{e}_3}{|\mathbf{w}+\mathbf{e}_3|} & \text{for } \mathbf{w} \neq -\mathbf{e}_3, \\ \begin{bmatrix} -\mathbf{e}_1, -\mathbf{e}_2, \mathbf{e}_3 \end{bmatrix} & \text{for } \mathbf{w} = -\mathbf{e}_3. \end{cases} \tag{4.10a}$$

Then, $\widetilde{\mathbf{H}}[\mathbf{w}]$ is orthonormal with $\widetilde{\mathbf{H}}[\mathbf{w}] = \widetilde{\mathbf{H}}[\mathbf{w}]^T = \widetilde{\mathbf{H}}[\mathbf{w}]^{-1}$ and maps $\mathbf{e}_3$ to $-\mathbf{w}$. Define

$$\mathbf{H}[\mathbf{w}] := \begin{bmatrix} \widetilde{\mathbf{H}}[\mathbf{w}]\mathbf{e}_1, \widetilde{\mathbf{H}}[\mathbf{w}]\mathbf{e}_2 \end{bmatrix} \in \mathbb{R}^{3\times2}, \tag{4.10b}$$

i.e., $\text{range}(\mathbf{H}[\mathbf{w}]) \perp \mathbf{w}$ and $\text{range}(\mathbf{H}[\mathbf{w}])$ mimics $\mathcal{K}_h[\boldsymbol{m}^n]$ nodewise. Moreover, for any orthogonal matrix $\mathbf{T} \in \mathbb{R}^{3\times3}$ with $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$, the matrix

$$\mathbf{T}\,\mathbf{H}[\mathbf{Tw}] \in \mathbb{R}^{3\times2} \quad \text{instead of} \quad \mathbf{H}[\mathbf{w}] \in \mathbb{R}^{3\times2}, \tag{4.10c}$$

also satisfies $\text{range}(\mathbf{TH}[\mathbf{Tw}]) \perp \mathbf{w}$. Hence, $\text{range}(\mathbf{TH}[\mathbf{Tw}])$ still mimics $\mathcal{K}_h[\boldsymbol{m}^n]$ nodewise. The following theorem provides a linear system in $\mathbb{R}^{2N} \cong \mathcal{K}_h[\boldsymbol{m}^n]$ for the solution to (4.5). The proof is postponed to Section 4.6.2 below.

Theorem 4.3.1. *Given $\boldsymbol{m}^n \in \mathcal{M}_h$, recall $\mathbf{A}[\boldsymbol{m}^n] \in \mathbb{R}^{3N\times3N}$ and $\mathbf{b}[\boldsymbol{m}^n] \in \mathbb{R}^{3N}$ from Section 4.2.4. Define the block-diagonal matrix*

$$\mathbf{Q}[\boldsymbol{m}^n] := \begin{pmatrix} \mathbf{TH}[\mathbf{T}\boldsymbol{m}^n(z_1)] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{TH}[\mathbf{T}\boldsymbol{m}^n(z_2)] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{TH}[\mathbf{T}\boldsymbol{m}^n(z_N)] \end{pmatrix} \in \mathbb{R}^{3N\times2N}, \quad (4.11)$$

where $\mathbf{H}(\cdot)$ is defined in (4.10). Then, the matrix $\mathbf{Q}[\boldsymbol{m}^n]^T\mathbf{A}[\boldsymbol{m}^n]\mathbf{Q}[\boldsymbol{m}^n] \in \mathbb{R}^{2N\times2N}$ is positive definite and, in particular, regular. Moreover, the unique solution $\mathbf{x} \in \mathbb{R}^{2N}$ of

$$\left(\mathbf{Q}[\boldsymbol{m}^n]^T\mathbf{A}[\boldsymbol{m}^n]\mathbf{Q}[\boldsymbol{m}^n] \right) \mathbf{x} = \mathbf{Q}[\boldsymbol{m}^n]^T\,\mathbf{b}[\boldsymbol{m}^n], \tag{4.12}$$

and the unique solution $\boldsymbol{v}^n \in \mathcal{K}_h[\boldsymbol{m}^n]$ of the variational formulation (4.5) satisfy

$$\boldsymbol{v}^n = \sum_{j=1}^{3N} \mathbf{v}_j\boldsymbol{\phi}_j \quad \text{with} \quad \mathbf{v} := \mathbf{Q}[\boldsymbol{m}^n]\,\mathbf{x} \in \mathbb{R}^{3N}. \tag{4.13}$$

Remark 4.3.2. *(i) For the validity of Theorem 4.3.1, it is only relevant that $\mathbf{H}[\boldsymbol{m}^n(z_i)]$ has orthonormal columns and that $\text{range}(\mathbf{H}[\boldsymbol{m}^n(z_i)]) \perp \boldsymbol{m}^n(z_i)$ for all $i = 1, \ldots, N$. Given $\mathbf{w} \in \mathbb{R}^3$ with $|\mathbf{w}| = 1$, alternative strategies from [Rug16, Lemma 6.1.2] are, e.g.,*

- *either to set $\mathbf{z} := \frac{\mathbf{w}+\sigma\mathbf{e}_3}{|\mathbf{w}+\sigma\mathbf{e}_3|}$, where $\sigma = \text{sign}(\mathbf{w}_3)$;*

- *or to use the transformation matrix of the rotation around the axis $\mathbf{e}_3 \times \mathbf{w}$ by an angle φ such that $\cos\varphi = \mathbf{w} \cdot \mathbf{e}_3$ and $\sin\varphi = |\mathbf{w} \times \mathbf{e}_3|$.*

(ii) The orthogonal matrix $\mathbf{T}$ may change with the time-step t_n.

4.4 Preconditioning

To solve the tangent space system (4.12), we choose a preconditioner $\mathbf{P} \in \mathbb{R}^{2N \times 2N}$ and employ the GMRES algorithm [SS86, Saa03] to the preconditioned system

$$\mathbf{PA_Q}[m]\mathbf{x} := \mathbf{PQ}[m]^T \mathbf{A}[m]\mathbf{Q}[m]\mathbf{x} = \mathbf{PQ}[m]^T \mathbf{b}[m] =: \mathbf{Pb_Q}[m], \qquad (4.14)$$

where $\boldsymbol{m} = \boldsymbol{m}^n \in \boldsymbol{\mathcal{M}}_h$. In the following sections, we discuss possible constructions of $\mathbf{P}$, based on the approximation of the inverse of the symmetric part of $\mathbf{A_Q}[\boldsymbol{m}]$ with α replaced by the parameter

$$\alpha_{\mathbf{P}} \geq \alpha > 0. \qquad (4.15)$$

In particular, this includes the case $\alpha_{\mathbf{P}} = \alpha$. Note that GMRES requires only the action of the preconditioner $\mathbf{P}$ on a vector. Moreover, recall that $\mathbf{Q}[\boldsymbol{m}]$ from (4.11) implicitly depends on the arbitrary but fixed matrix $\mathbf{T} \in \mathbb{R}^{3 \times 3}$ from (4.10c). We refer to Section 4.4.2 below for the possible construction of the matrix $\mathbf{T}$, for given $\boldsymbol{m} \in \boldsymbol{\mathcal{M}}_h$.

4.4.1 Practical preconditioner

For general problems of type (4.12), the work [NS96] proposes (without a proof) to consider the practical preconditioner

$$\mathbf{P_Q}[\boldsymbol{m}] := \mathbf{Q}[\boldsymbol{m}]^T \big(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\text{ex}}^2 k\,\mathbf{L}\big)^{-1}\mathbf{Q}[\boldsymbol{m}] \in \mathbb{R}^{2N \times 2N}. \qquad (4.16)$$

The following theorem discusses the performance of GMRES with the preconditioner $\mathbf{P_Q}[\boldsymbol{m}]$. Its proof is postponed to Section 4.6.6.

Theorem 4.4.1. *Let $\alpha_{\mathbf{P}} \geq \alpha$. Consider the preconditioned GMRES algorithm with the preconditioner $\mathbf{P_Q}[\boldsymbol{m}]$ from (4.16) for the solution of (4.14) with the initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^{2N}$. For $\ell \in \mathbb{N}_0$, let $\mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}$ denote the GMRES iterates with the corresponding residuals*

$$\mathbf{r}^{(\ell)} := \mathbf{P_Q}[\boldsymbol{m}]\mathbf{b_Q}[\boldsymbol{m}] - \mathbf{P_Q}[\boldsymbol{m}]\mathbf{A_Q}[\boldsymbol{m}]\,\mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}.$$

Then, there exists $0 < \kappa < 1$ such that

$$\|\mathbf{r}^{(\ell)}\|_{\boldsymbol{m}} \leq (1 - \kappa)^{\ell/2}\,\|\mathbf{r}^{(0)}\|_{\boldsymbol{m}} \quad \textit{for all } \ell \in \mathbb{N}, \qquad (4.17)$$

where $\|\cdot\|_{\boldsymbol{m}}$ is the norm induced by the energy scalar product

$$\langle\!\langle\!\langle \mathbf{x}\,,\,\mathbf{y} \rangle\!\rangle\!\rangle_{\boldsymbol{m}} := \mathbf{x} \cdot \big(\mathbf{P_Q}[\boldsymbol{m}]\big)^{-1}\mathbf{y} \quad \textit{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}. \qquad (4.18)$$

Moreover, there hold the following assertions (i)–(ii), where $C > 1$ depends only on $C_{\mathbf{mesh}}$:
(i) In general, κ can be chosen such that

$$\kappa \geq \left[C \left(\frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha} \right)^2 \left(1 + \frac{\ell_{\text{ex}}^2 k}{\alpha_{\mathbf{P}} h^2} \right) \right]^{-1} > 0\,. \qquad (4.19)$$

(ii) If $\mathbf{T}$ is chosen such that $1 + (\mathbf{T}\boldsymbol{m}(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, then κ can be chosen such that

$$\kappa \geq \left[C \left(\frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha} \right)^2 \left(1 + \frac{\ell_{\text{ex}}^2 k}{\alpha_{\mathbf{P}} \gamma^4} \|\nabla \boldsymbol{m}\|_{\boldsymbol{L}^\infty(\Omega)}^2 \right) \right]^{-1} > 0\,. \qquad (4.20)$$

Remark 4.4.2. *(i) The bound on $0 < \kappa < 1$ from Theorem 4.4.1(i) requires a CFL-type coupling $k = \mathcal{O}(h^2)$ to prevent deterioration $\kappa \to 0$ as $h \to 0$.*

(ii) The additional assumption of Theorem 4.4.1(ii) can always be achieved by an appropriate choice of $\mathbf{T}$. Then, $0 < \kappa < 1$ is bounded away from 0 as long as no finite-time blow-up occurs. Moreover, as can be expected for parabolic problems, the bound improves as $k \to 0$.

(iii) The preconditioner $\mathbf{P_Q}[\boldsymbol{m}]$ involves the inversion of the symmetric part $\left(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\mathrm{ex}}^2 k\mathbf{L}\right)$. In practice, this can be done by a multigrid method. We note that the symmetric part is, in particular, independent of $\boldsymbol{m}$ (i.e., of the time-step).

4.4.2 Practical computation of T

Theorem 4.4.1(ii) requires that

$$1 + \boldsymbol{m}(z) \cdot \mathbf{T}\mathbf{e}_3 \geq \gamma > 0 \quad \text{for all nodes } z \in \mathcal{N}_h. \tag{4.21}$$

For given $\boldsymbol{m} \in \boldsymbol{\mathcal{M}}_h$ and in view of (4.20), we thus aim to choose the matrix $\mathbf{T} \in \mathbb{R}^{3\times3}$ from (4.10c) such that γ in (4.21) can be chosen as large as possible. To this end, define $d_\ell^\star \in [0, 2]$ by

$$d_\ell^+ := 1 - \max_{z \in \mathcal{N}_h} (\boldsymbol{m}(z))_\ell, \quad d_\ell^- := 1 + \min_{z \in \mathcal{N}_h} (\boldsymbol{m}(z))_\ell, \quad \text{for all } \ell \in \{1, 2, 3\}. \tag{4.22}$$

In addition, define $\mathbf{T}_\ell^\star \in \mathbb{R}^{3\times3}$ with $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$ by

$$\mathbf{T}_1^+ := [-\mathbf{e}_3, \mathbf{e}_2, -\mathbf{e}_1], \quad \mathbf{T}_1^- := [\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1], \tag{4.23a}$$

$$\mathbf{T}_2^+ := [\mathbf{e}_1, -\mathbf{e}_3, -\mathbf{e}_2], \quad \mathbf{T}_2^- := [\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2], \tag{4.23b}$$

$$\mathbf{T}_3^+ := [\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_3], \quad \mathbf{T}_3^- := [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]. \tag{4.23c}$$

For all $\ell \in \{1, 2, 3\}$ and $\star \in \{+, -\}$, it holds that

$$1 + \boldsymbol{m}(z) \cdot \mathbf{T}_\ell^\star \mathbf{e}_3 \geq d_\ell^\star \in [0, 2] \quad \text{for all nodes } z \in \mathcal{N}_h. \tag{4.24}$$

Hence, (4.21) holds with $\gamma \in [0, 2]$ being the maximum $d_\ell^\star$, and $\mathbf{T} = \mathbf{T}_\ell^\star$ in (4.23). We note that other (more sophisticated) strategies are possible. Finally, we note that our (very simple) construction leads to $\gamma = 0$, if and only if

$$\{\pm\mathbf{e}_1, \pm\mathbf{e}_2, \pm\mathbf{e}_3\} \subseteq \{\boldsymbol{m}(z) : z \in \mathcal{N}_h\}.$$

4.4.3 Stationary preconditioner

We consider a preconditioner which is independent of the time-step. Define the 2D-equivalent to the basis from (4.7) by

$$\boldsymbol{\psi}_{2(j-1)+\ell} := \varphi_j \, \boldsymbol{e}_\ell : \Omega \to \mathbb{R}^2 \quad \text{for all } j = 1, \ldots, N \text{ and all } \ell = 1, 2. \tag{4.25}$$

Similarly to $\mathbf{M}, \mathbf{L} \in \mathbb{R}^{3N\times3N}$ from Section 4.2.4, define the matrices $\mathbf{M}_{\mathrm{2D}}, \mathbf{L}_{\mathrm{2D}} \in \mathbb{R}^{2N\times2N}$, which correspond to the nodal basis $(\boldsymbol{\psi}_i)_{i=1}^{2N}$ of $(\mathcal{S}_h)^2$ from (4.25):

- $(\mathbf{M}_{2\mathrm{D}})_{ij} := \int_\Omega \boldsymbol{\psi}_j \cdot \boldsymbol{\psi}_i \, dx$ is the (symmetric and positive definite) mass matrix,

- $(\mathbf{L}_{2\mathrm{D}})_{ij} := \int_\Omega \nabla \boldsymbol{\psi}_j \cdot \nabla \boldsymbol{\psi}_i \, dx$ is the (symmetric and positive semidefinite) stiffness matrix.

Then, consider the stationary preconditioner

$$\mathbf{P}_{2\mathrm{D}} := \left(\alpha_{\mathbf{P}} \mathbf{M}_{2\mathrm{D}} + \ell_{\mathrm{ex}}^2 k \, \mathbf{L}_{2\mathrm{D}} \right)^{-1} \in \mathbb{R}^{2N \times 2N}. \tag{4.26}$$

The following theorem discusses the performance of GMRES with the preconditioner $\mathbf{P}_{2\mathrm{D}}$; see Remark 4.4.2(ii) for some comments on the result. The proof is postponed to Section 4.6.5.

Theorem 4.4.3. *Let $\alpha_{\mathbf{P}} \geq \alpha$. Consider the preconditioned GMRES algorithm with $\mathbf{P}_{2\mathrm{D}}$ from (4.26) for the solution of (4.14) with the initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^{2N}$. For $\ell \in \mathbb{N}_0$, let $\mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}$ denote the GMRES iterates with the corresponding residuals*

$$\mathbf{r}^{(\ell)} := \mathbf{P}_{2\mathrm{D}} \mathbf{b}_{\mathbf{Q}}[\boldsymbol{m}] - \mathbf{P}_{2\mathrm{D}} \mathbf{A}_{\mathbf{Q}}[\boldsymbol{m}] \, \mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}.$$

If $\mathbf{T}$ is chosen such that $1 + (\mathbf{T}\boldsymbol{m}(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, then there exists $0 < \kappa < 1$ such that

$$\|\mathbf{r}^{(\ell)}\| \leq (1 - \kappa)^{\ell/2} \, \|\mathbf{r}^{(0)}\| \quad \text{for all } \ell \in \mathbb{N}, \tag{4.27}$$

where $\|\cdot\|$ denotes the norm induced by the corresponding energy scalar product

$$\langle\!\langle\!\langle \mathbf{x}, \mathbf{y} \rangle\!\rangle\!\rangle := \mathbf{x} \cdot \mathbf{P}_{2D}^{-1} \mathbf{y} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}.$$

Moreover, with $C > 1$ depending only on $C_{\mathbf{mesh}}$, κ can be chosen such that

$$\kappa \geq \left[C \left(\frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha} \right)^2 \left(1 + \gamma^{-2} + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_{\mathbf{P}} \gamma^6} \|\nabla \boldsymbol{m}\|_{\boldsymbol{L}^\infty(\Omega)}^2 \right)^4 \right]^{-1} > 0. \tag{4.28}$$

4.4.4 Jacobi-type preconditioner

Consider the following approximation to the stationary preconditioner $\mathbf{P}_{2\mathrm{D}}$ from (4.26): Recalling $M_{ij}, L_{ij} \in \mathbb{R}$ from (4.8a), we set

$$\mathbf{P}_i^{\mathrm{jac}} := \left(\alpha_{\mathbf{P}} M_{ii} + \ell_{\mathrm{ex}}^2 k L_{ii} \right)^{-1} \mathbf{I}_{2 \times 2} \in \mathbb{R}^{2 \times 2} \quad \text{for all } i = 1, \dots, N$$

and define the stationary Jacobi-type preconditioner

$$\mathbf{P}_{2\mathrm{D}}^{\mathrm{jac}} := \begin{pmatrix} \mathbf{P}_1^{\mathrm{jac}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2^{\mathrm{jac}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{P}_N^{\mathrm{jac}} \end{pmatrix} \in \mathbb{R}^{2N \times 2N}.$$

Given $\boldsymbol{m} \in \mathcal{M}_h$, the matrix

$$\mathbf{Q}[\boldsymbol{m}]^T \mathbf{P}_{3\mathrm{D}}^{\mathrm{jac}} \mathbf{Q}[\boldsymbol{m}] \in \mathbb{R}^{2N \times 2N}, \quad \text{where } \left[\mathbf{P}_{3\mathrm{D}}^{\mathrm{jac}} \right]_{ij} := \begin{cases} \left[(\alpha_{\mathbf{P}} \mathbf{M} + \ell_{\mathrm{ex}}^2 k \, \mathbf{L})^{-1} \right]_{ii} & \text{for } i = j, \\ 0 & \text{else,} \end{cases}$$

is the Jacobi-type approximation of $\mathbf{P}_{\mathbf{Q}}[\boldsymbol{m}]$ from (4.16). The following proposition states that the definitions of the Jacobi-type preconditioners coincide.

Proposition 4.4.4. *It holds that* $\mathbf{P}^{\text{jac}} := \mathbf{P}^{\text{jac}}_{\text{2D}} = \mathbf{Q}[\boldsymbol{m}]^T \mathbf{P}^{\text{jac}}_{\text{3D}} \mathbf{Q}[\boldsymbol{m}]$.

Proof. The statement follows from the representations of $\mathbf{M}$ and $\mathbf{L}$ given in (4.8) as well as the block-diagonal definition of $\mathbf{Q}[\boldsymbol{m}]$, where the blocks $\mathbf{TH}[\mathbf{T}\boldsymbol{m}(z_i)]$ have orthonormal columns for all nodes $z_i \in \mathcal{N}_h$. $\qquad\square$

4.5 Numerics

In this section, we underpin our theoretical findings with numerical experiments. To this end, we employ our Python code for computational micromagnetics Commics [Pfe, PRS$^+$20], which is based on the open-source multiphysics finite element library Netgen/NG-Solve [ngs]. The computation of the stray field is based on the hybrid FEM-BEM method of [FK90], which requires the evaluation of the double-layer integral operator associated with the Laplace equation. This part of the code exploits the open-source Galerkin boundary element library BEM++ [ŚBA$^+$15]. Moreover, we couple Netgen/NGSolve and BEM++ with ngbem [Rie].

To solve the (preconditioned) linear system (4.12), we employ the (preconditioned) GMRES algorithm [SS86, Saa03]. Our implementation uses the GMRES implementation from SciPy [JOP$^+$], where we employ the iteration tolerance $\varepsilon = 10^{-8}$. To save memory, GMRES is restarted every 20 iterations. Note that this is commonly referred to as restarted GMRES; cf., e.g., [Saa03, Algorithm 6.11]. As initial value for the GMRES iteration, we always choose $\mathbf{x}_0 = \mathbf{0}$.

The experiments of this section are focused on the number of iterations of the preconditioned GMRES algorithm. As suggested in Remark 4.4.2(iii), to apply the inverse matrices appearing in the preconditioners from Section 4.4 (i.e., in (4.16) and (4.26)) to a vector, we use the classical (Ruge–Stüben) algebraic multigrid method implemented in PyAMG [OS18].

Remark 4.5.1. *The 'inverse' matrices appearing in the practical preconditioner $\mathbf{P_Q}[\cdot]$ from (4.16) and in the stationary preconditioner $\mathbf{P}_{\text{2D}}$ from (4.26) are block-diagonal matrices and consist of the same $N \times N$-stationary matrix block. Moreover, $\mathbf{Q}[\cdot]$ from (4.11) in $\mathbf{P_Q}[\cdot]$ has a block-diagonal form and is explicitly available at each time-step. Hence, $\mathbf{P}_{\text{2D}}$ and $\mathbf{P_Q}[\cdot]$ have similar computational complexity. The Jacobi-type preconditioner $\mathbf{P}^{\text{jac}}$ from Proposition 4.4.4 is computationally favorable as its construction only involves the inversion of diagonal matrices.*

4.5.1 An academic example

We investigate the dependence of the total number of GMRES iterations on the mesh-size h for the proposed preconditioners. To this end, we adapt the setting of [PRS18b, Section 6.1]: On $\Omega := (0,1)^3$, we employ the initial value and applied field

$$\boldsymbol{m}^0 := (1,0,0) \quad \text{and} \quad \boldsymbol{f}(x_1, x_2, x_3) := 10\,(\sin(x_1), \cos(x_1), 0),$$

respectively. We set $T = 1$, $\alpha = 0.5$, $\ell^2_{\text{ex}} = 10$. The only lower-order contribution in $\boldsymbol{\pi}(\boldsymbol{m})$ is the stray field. With this setting, the magnetization is expected to align itself in the direction of the applied field $\boldsymbol{f}$. For the time discretization, we fix $k = 1/100$. For the

space discretization, we consider the following structured meshes generated by Commics: For $j \in \mathbb{N}$ the domain Ω is decomposed into j^3 cubic cells of edge length $1/j$. Each of these cells is divided into six equally sized tetrahedra with diameter $\sqrt{3}/j$. This yields meshes consisting of $(j+1)^3$ vertices and $6j^3$ elements. The experiment is performed for $j \in \{10, 11, \ldots, 19\} \cup \{20, 25, \ldots, 50\}$. We compare the performance of the preconditioners proposed in Section 4.4 with $\alpha_{\mathbf{P}} = 1$ for the iterative solution of the underlying linear system (4.12). Moreover, we always use $\mathbf{T} := \mathbf{I}_{3 \times 3}$, i.e., we always employ the standard choice (4.10b).
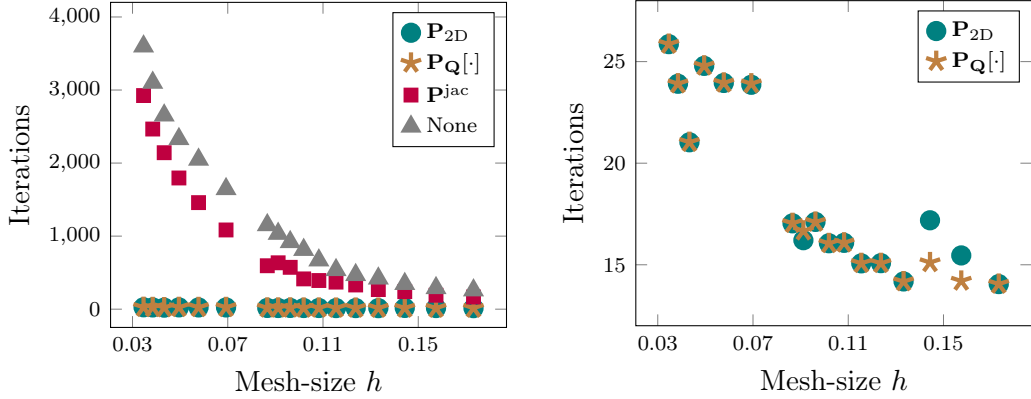

Figure 4.1: Experiment of Section 4.5.1: Average number of GMRES iterations for different mesh-sizes. The right plot is a zoom of the left plot with the results for $\mathbf{P}_{2D}$ and $\mathbf{P}_{\mathbf{Q}}[\cdot]$.

In Figure 4.1, we plot the average number of GMRES iterations for the different meshes. As expected, no preconditioning requires the most iterations. The Jacobi-type preconditioner $\mathbf{P}^{\mathrm{jac}}$ brings a slight improvement. However, both the results obtained with no preconditioning and $\mathbf{P}^{\mathrm{jac}}$ are not robust with respect to the mesh-size h. In contrast to that, the stationary preconditioner $\mathbf{P}_{2D}$ and the practical preconditioner $\mathbf{P}_{\mathbf{Q}}[\cdot]$ require significantly less iterations. Moreover, these two options are robust with respect to the mesh-size h. In conclusion, this experiment suggests the use of either the stationary preconditioner $\mathbf{P}_{2D}$ or the practical preconditioner $\mathbf{P}_{\mathbf{Q}}[\cdot]$.

4.5.2 μMAG standard problem #4

We investigate the practical applicability of the preconditioners proposed in Section 4.4 by computing a physically relevant example. To this end, we consider μMAG standard problem #4 [MUM], which simulates the switching of the magnetization in a thin permalloy layer.

The domain Ω is a rectangular cuboid of length 500 nm, width 125 nm, and thickness 3 nm. The involved physical constants and material parameters are the gyromagnetic ratio $\gamma_0 = 2.211 \cdot 10^5$ m/C, the permeability of vacuum $\mu_0 = 4\pi \cdot 10^{-7}$ N/A^2, the saturation magnetization $M_{\mathrm{s}} = 8.0 \cdot 10^5$ A/m, the exchange stiffness constant $A = 1.3 \cdot 10^{-11}$ J/m, and the Gilbert damping constant $\alpha = 0.02$. Starting from a so-called S-state [MUM], the experiment consists in applying the constant applied field $\mu_0 \boldsymbol{H}_{\mathrm{ext}} = (-35.5, -6.3, 0)$ mT for 3 ns.

For the rescaled form (4.1) of LLG, the above physical quantities lead to the parameters $\ell_{\mathrm{ex}} = \sqrt{2A/(\mu_0 M_{\mathrm{s}}^2)}$, $T = 3 \cdot 10^{-9} \gamma_0 M_{\mathrm{s}}$, and $\boldsymbol{f} = \boldsymbol{H}_{\mathrm{ext}}/M_{\mathrm{s}}$, and $\boldsymbol{\pi}(\boldsymbol{m})$ includes only the stray field.

For the space discretization, we employ a fixed structured mesh with constant mesh-size $h = 4.2\,\mathrm{nm}$ consisting of $97\,200$ elements and $24\,978$ vertices. For the time discretization, we will consider the physical time-step size Δt, which is connected to the rescaled time-step size k via the relation $k = \gamma_0 M_{\mathrm{s}} \Delta t$.

In the following subsections, we numerically investigate the performance of the preconditioners proposed in Section 4.4 for the simulation of μMAG standard problem #4. In particular, we study the impact of the choices of $\alpha_{\mathbf{P}}$ (Section 4.5.2.1) and the matrix $\mathbf{T}$ (Section 4.5.2.2) on the performance as well as the time-step size robustness of the preconditioners (Section 4.5.2.3). If not stated otherwise, we use the time-step size $1/8\,\mathrm{ps}$, $\alpha_{\mathbf{P}} = 1$ and the adaptive strategy from Section 4.4.2 for $\mathbf{T}$.

4.5.2.1 The impact of $\alpha_{\mathbf{P}}$

In this experiment, we numerically investigate the optimal choice of $\alpha_{\mathbf{P}}$ in (4.15). In Figure 4.2, we plot the required number of GMRES iterations for different preconditioners over time.

(a) $\alpha_{\mathbf{P}} = 0.02 = \alpha$.

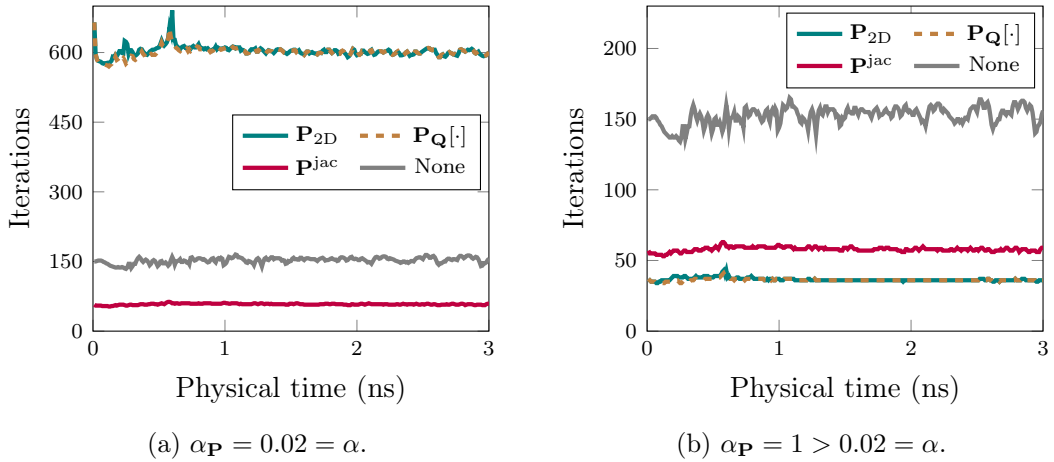(b) $\alpha_{\mathbf{P}} = 1 > 0.02 = \alpha$.

Figure 4.2: Experiment of Section 4.5.2: GMRES iterations over time.

To start with, we consider the choice $\alpha_{\mathbf{P}} = \alpha = 0.02$ (see Figure 4.2a for the corresponding results). The Jacobi-type preconditioner $\mathbf{P}^{\mathrm{jac}}$ requires less iterations than no preconditioning. The stationary preconditioner $\mathbf{P}_{2\mathrm{D}}$ and the practical preconditioner $\mathbf{P}_{\mathbf{Q}}[\cdot])$ fail completely and require significantly more GMRES iterations than no preconditioning. This might be the effect of the skew-symmetric part $\mathbf{S}[\cdot]$ of the (unconstrained system) matrix $\mathbf{A}[\cdot]$ from (4.9). Here, the scaling of $\mathbf{S}[\cdot]$ is similar to the scaling of the mass matrix, but unlike $\mathbf{M}$, it lacks the factor α.

We repeat the experiment, but we choose $\alpha_{\mathbf{P}} = 1 > 0.02 = \alpha$ (see Figure 4.2b). No preconditioning requires the most iterations. With this choice of $\alpha_{\mathbf{P}}$, using one of the proposed preconditioners significantly reduces the number of iterations. Here, $\mathbf{P}_{2\mathrm{D}}$ and $\mathbf{P}_{\mathbf{Q}}[\cdot]$

require the fewest iterations. This experiment indicates that the stationary preconditioner $\mathbf{P}_{2D}$ and the practical preconditioner $\mathbf{P_Q}[\cdot]$ with $\alpha_\mathbf{P} = 1$ lead to a satisfactory performance.
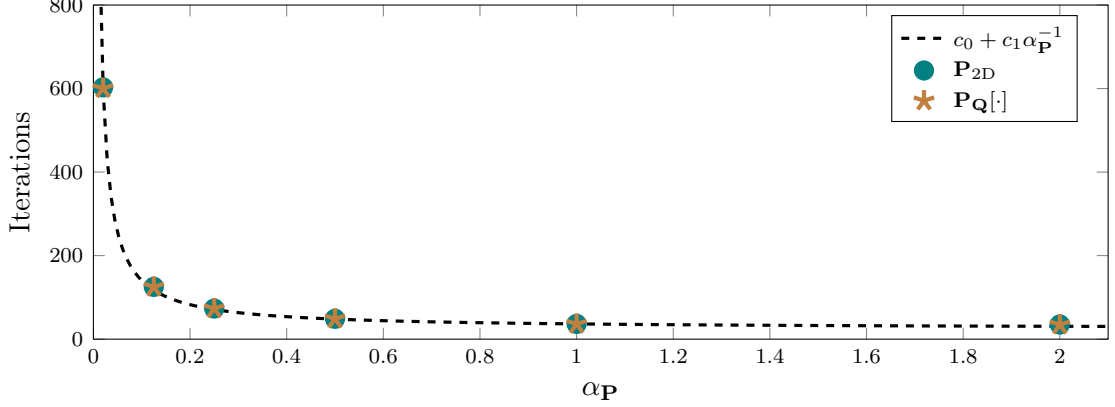

Figure 4.3: Experiment of Section 4.5.2.1: Average number of GMRES iterations. The values c_0 and c_1 for the black dashed line are obtained by fitting to the iteration numbers of $\alpha_\mathbf{P} = 0.02$ and $\alpha_\mathbf{P} = 1$.

To empirically determine the optimal choice for $\alpha_\mathbf{P}$, we consider the preconditioners $\mathbf{P}_{2D}$ and $\mathbf{P_Q}[\cdot]$ and vary $\alpha_\mathbf{P}$. In Figure 4.3, we plot the resulting averaged number of the required GMRES iterations as a function of $\alpha_\mathbf{P}$. Values of $\alpha_\mathbf{P}$ larger than the damping parameter α result in significantly less iterations: Empirically, the number of GMRES iterations seems to qualitatively behave like $c_0 + c_1/\alpha_\mathbf{P}$ for some $c_0, c_1 \in \mathbb{R}$. However, choosing $\alpha_\mathbf{P}$ larger than 1 seems to have little or no effect. In conclusion, we suggest choosing $\alpha_\mathbf{P} = 1$.

4.5.2.2 Adaptive vs. fixed **T**

In this experiment, we discuss the impact of the adaptive strategy for $\mathbf{T} \in \mathbb{R}^{3\times3}$ described in Section 4.4.2. In Figure 4.4, we plot the evolution of $d_\ell^\star \in [0, 2]$ from (4.22), where $\ell \in \{1, 2, 3\}$ and $\star \in \{+, -\}$ and $d_{\mathrm{adapt}} := \max_{\ell\in\{1,2,3\}}\{d_\ell^+, d_\ell^-\}$. Recall from (4.24) that we can choose $\gamma = d_{\mathrm{adapt}}$ for adaptive $\mathbf{T}$ and $\gamma = d_\ell^\star$ for fixed $\mathbf{T} := \mathbf{T}_\ell^\star$. For adaptive $\mathbf{T}$, in our example we always have that $1 + (\mathbf{T}_n \boldsymbol{m}_h^n(z))_3 \geq \gamma > 0$, i.e., Theorem 4.4.1(ii) and Theorem 4.4.3 apply.

In Figure 4.5a, we consider the stationary preconditioner $\mathbf{P}_{2D}$. We plot the evolution of the GMRES iteration numbers for the adaptive strategy and for the fixed choices $\mathbf{T} := \mathbf{T}_\ell^\star$ from (4.23), where $\ell \in \{1, 2, 3\}$ and $\star \in \{+, -\}$. In Figure 4.5b, we repeat this experiment for the practical preconditioner $\mathbf{P_Q}[\cdot]$.

As can be seen in Figure 4.5a, an adaptive choice of $\mathbf{T}$ is not always the perfect choice, however, it avoids the increased iteration number that can be observed in many cases for fixed $\mathbf{T}$. Yet, for the relation for the iteration number of fixed $\mathbf{T} = \mathbf{T}_\ell^\star$ and the corresponding $d_\ell^\star$, the picture is not complete: In Figure 4.5b, all options appear to be equally good, even though, e.g., for fixed $\mathbf{T} := \mathbf{T}_1^-$, it holds that $d_1^- \approx 0$ most of the time (see Figure 4.4).

In conclusion, our experiment suggests to use the adaptive choice for $\mathbf{T}$. However, a full understanding of the effect of the choice of $\mathbf{T}$ will require further investigations.
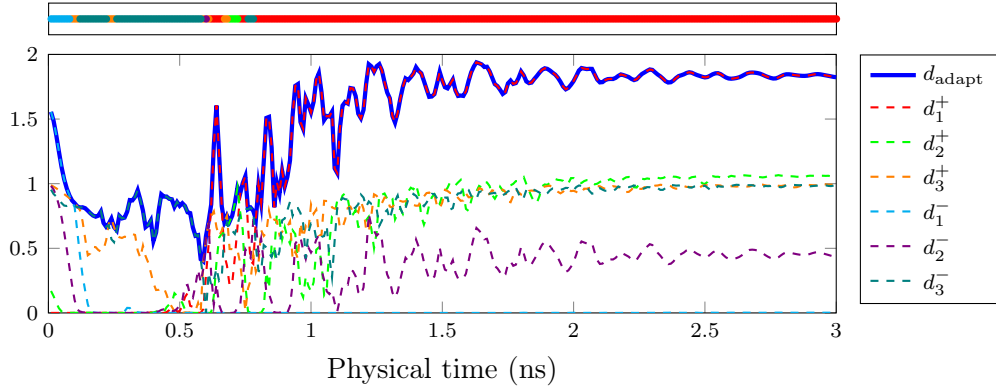
Figure 4.4: Experiment of Section 4.5.2.2: Evolution of $d_\ell^\star$ with $\ell \in \{1,2,3\}$ and $\star \in \{+,-\}$, and $d_{\text{adapt}} := \max_{\ell \in \{1,2,3\}} \{d_\ell^+, d_\ell^-\}$ over time. The above line indicates the current state of d_{adapt}.

$\Delta t =$	$1/4\,\text{ps}$	$1/8\,\text{ps}$	$1/16\,\text{ps}$	$1/32\,\text{ps}$
$\mathbf{P}_{2\text{D}}$	39.03	36.55	34.71	33.38
$\mathbf{P_Q}[\cdot]$	38.92	36.16	33.96	32.69

Table 4.1: Experiment of Section 4.5.2.3: Average number of GMRES iterations for different time-step sizes Δt.

4.5.2.3 Time-step robustness

In this experiment, we test the robustness of the preconditioners with respect to different time-step sizes. For the simulation we vary the time-step size and choose $\Delta t = 1/4\,\text{ps}, 1/8\,\text{ps}, 1/16\,\text{ps}$ and $1/32\,\text{ps}$.

In Table 4.1, the average number of GMRES iterations for the stationary preconditioner $\mathbf{P}_{2\text{D}}$ and the practical preconditioner $\mathbf{P_Q}[\cdot]$ are compared. As expected, the iteration numbers improve slightly for smaller time-step sizes; see Remark 4.4.2(ii). While the time-step sizes vary up to the multiplicative factor 8, the number of iterations changes only by a few percent. Hence, this experiment indicates that the preconditioners might not only be robust with respect to the mesh-size (see Section 4.5.1), but also robust with respect to the time-step size.

4.5.3 μMAG standard problem #5

We consider μMAG standard problem #5 [MUM] and repeat the experiments presented in Section 4.5.2 (μMAG standard problem #4), as well as the experiment from Section 4.5.1 (academic setting).

The domain Ω is a rectangular cuboid of length $100\,\text{nm}$, width $100\,\text{nm}$, and thickness $10\,\text{nm}$, aligned with the x, y, and z axes of a Cartesian coordinate system, with origin at the center of the cuboid. The material parameters coincide with those used in Section 4.5.2, except for the Gilbert damping parameter ($\alpha = 0.1$ here). The experiment consists in
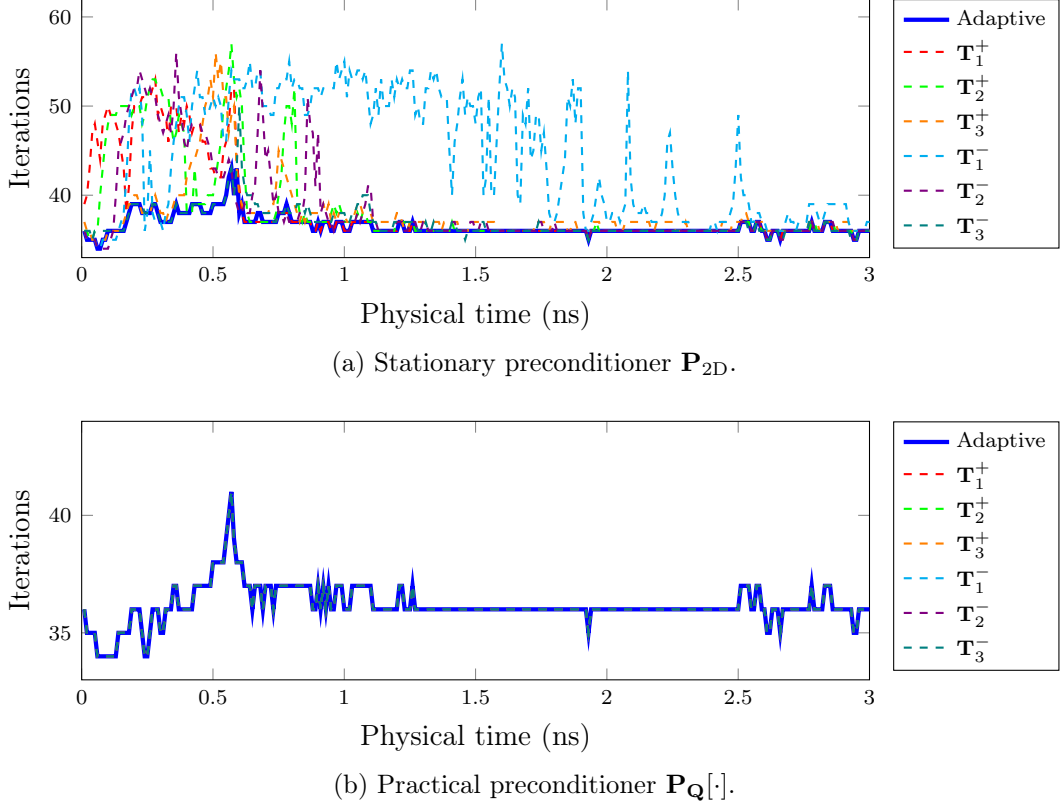
(a) Stationary preconditioner $\mathbf{P}_{\mathrm{2D}}$.

(b) Practical preconditioner $\mathbf{P}_{\mathbf{Q}}[\cdot]$.

Figure 4.5: Experiment of Section 4.5.2.2: GMRES iterations for the preconditioners over time.

simulating for $6\,\mathrm{ns}$ the movement of a magnetic vortex in response to a constant electric current flowing in the plane of the sample. The initial state is obtained by solving (4.1) for $\boldsymbol{f} = \boldsymbol{0}$, $\boldsymbol{\pi}(\boldsymbol{m})$ consisting of only the stray field, and the initial condition $\boldsymbol{m}^0(x, y, z) = (-y, x, R)/\sqrt{x^2 + y^2 + R^2}$ with $R = 10\,\mathrm{nm}$ for a sufficiently long time, until the system relaxes to equilibrium. To simulate the current-driven movement of the vortex, we solve (4.1) with this relaxed magnetization configuration as initial condition, $T = 6 \cdot 10^{-9} \gamma_0 M_{\mathrm{s}}$, and $\boldsymbol{\pi}(\boldsymbol{m})$ consisting of the stray field and the Zhang–Li spin-torque term [ZL04, TNMS05], which takes the expression $\boldsymbol{m} \times (\boldsymbol{u} \cdot \nabla)\boldsymbol{m} + \beta(\boldsymbol{u} \cdot \nabla)\boldsymbol{m}$, with $\boldsymbol{u} = (72.17, 0, 0)/(\gamma_0 M_{\mathrm{s}})$ and $\beta = 0.05$.

To repeat the experiments from Section 4.5.2 for this problem setting, we employ a fixed structured mesh with constant mesh-size $h = 3.46\,\mathrm{nm}$ consisting of $75\,000$ elements and $15\,606$ vertices for the space discretization. For the time discretization, we consider the physical time-step size $\Delta t = 1/8\,\mathrm{ps}$. The results qualitatively agree with those of Section 4.5.2. For a snapshot of the results, we refer to Figure 4.6 and Figure 4.7.

To investigate the robustness of the preconditioners with respect to varying mesh-sizes, we choose $\alpha_{\mathbf{P}} = 1$, $\mathbf{T} \equiv \mathbf{T}_3^-$ (coincides with adaptive choice; cf. Figure 4.7) and repeat the experiment with constant mesh-sizes $h = 2.17\,\mathrm{nm}, 1.73\,\mathrm{nm}, 1.44\,\mathrm{nm}, 1.15\,\mathrm{nm}$ and $h = 0.87\,\mathrm{nm}$. The results shown in Figure 4.8 confirm the observations from Section 4.5.1 also

for μMAG standard problem #5.

(a) $\alpha_{\mathbf{P}} = 0.1 = \alpha$.

(b) $\alpha_{\mathbf{P}} = 1 > 0.1 = \alpha$.
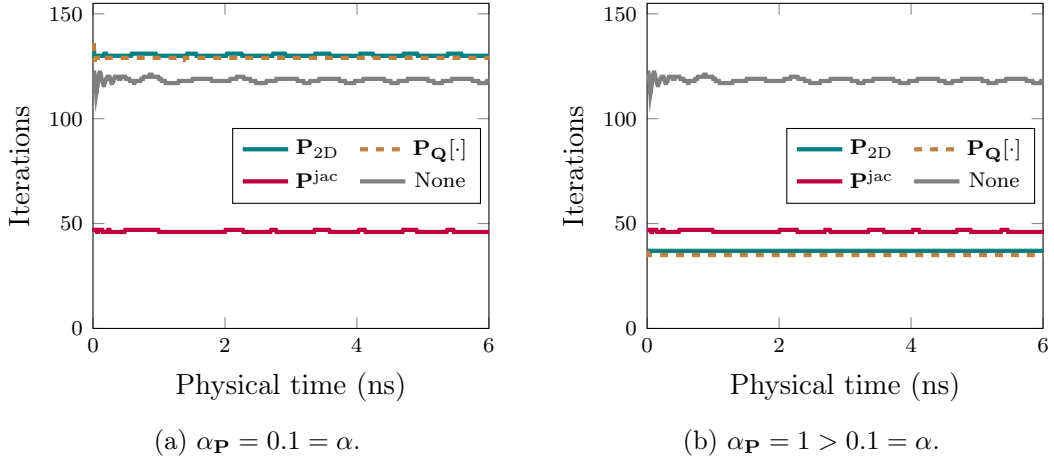
Figure 4.6: Experiment of Section 4.5.3: GMRES iterations over time.
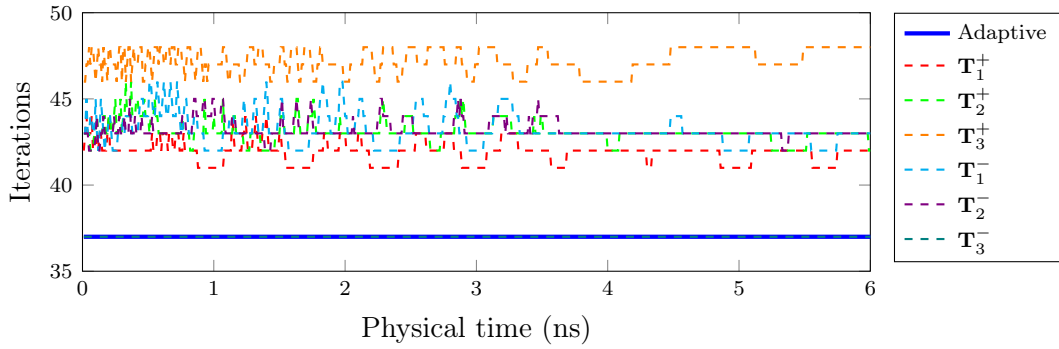
Figure 4.7: Experiment of Section 4.5.3: GMRES iterations for the stationary precondi-
tioner $\mathbf{P}_{2\mathrm{D}}$ over time ($\alpha_{\mathbf{P}} = 1$). The adaptive strategy coincides with $\mathbf{T}_3^-$.

4.6 Proof of main results

For a matrix $\mathbf{H} \in \mathbb{R}^{d \times d}$, let $|\mathbf{H}| := \sup_{\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}} |\mathbf{A}\mathbf{x}|/|\mathbf{x}|$ be the matrix norm induced by
the Euclidean norm $|\mathbf{x}|$ on $\mathbb{R}^d$.

To abbreviate notation, we write $\langle \Omega, \cdot \rangle \cdot$ for all L^2-scalar products (including vector-valued
spaces) and $\|\cdot\|_\Omega = \|\cdot\|_{\boldsymbol{L}^2(\Omega)}$ for the corresponding L^2-norm. The L^∞-norm is abbreviated
by $\|\cdot\|_\infty = \|\cdot\|_{\boldsymbol{L}^\infty(\Omega)}$.

For $a, b \in \mathbb{R}^+$ with $a \leq Cb$, we write $a \lesssim b$, if the constant $C > 0$ is independent of the
discretization parameters and clear from the context. Finally, $a \simeq b$ abbreviates $a \lesssim b \lesssim a$.

Figure 4.8: Experiment of Section 4.5.3: Average number of GMRES iterations for different mesh-sizes. The right plot is a zoom of the left plot with the results for $\mathbf{P}_{2\mathrm{D}}$ and $\mathbf{P}_{\mathbf{Q}}[\cdot]$.

4.6.1 Auxiliary mappings

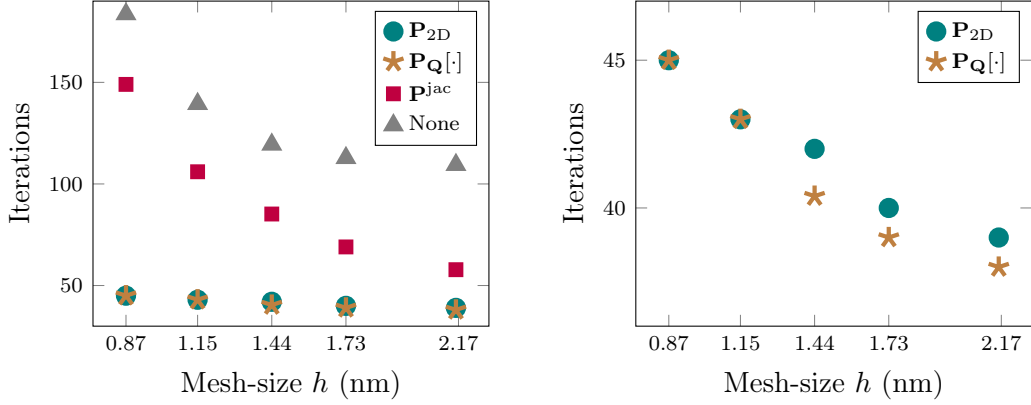Recall the hat functions $(\varphi_i)_{i=1}^N$, the nodal basis $(\boldsymbol{\psi}_i)_{i=1}^{2N}$ of $(\mathcal{S}_h)^2$ from (4.25), and $(\boldsymbol{\phi}_i)_{i=1}^{3N}$ of $\boldsymbol{\mathcal{S}}_h = (\mathcal{S}_h)^3$ from (4.7). Recall the definitions of $\mathbf{H}[\cdot] \in \mathbb{R}^{3\times 2}$ and $\mathbf{T} \in \mathbb{R}^{3\times 3}$ from (4.10) and $\mathbf{Q}[\cdot] \in \mathbb{R}^{3N\times 2N}$ from (4.11). Given $\boldsymbol{\mu} \in \mathcal{M}_h$, define the mappings

$$\mathbb{P}_h[\boldsymbol{\mu}] : (\mathcal{S}_h)^2 \to \mathcal{K}_h[\boldsymbol{\mu}] \subsetneq \boldsymbol{\mathcal{S}}_h : \boldsymbol{w} \mapsto \sum_{i=1}^N \left(\mathbf{TH}[\mathbf{T}\boldsymbol{\mu}(z_i)]\boldsymbol{w}(z_i) \right) \varphi_i, \tag{4.29a}$$

$$\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}] : \mathbb{R}^{2N} \to \mathcal{K}_h[\boldsymbol{\mu}] \subsetneq \boldsymbol{\mathcal{S}}_h : \mathbf{x} \mapsto \sum_{i=1}^{3N} \left(\mathbf{Q}[\boldsymbol{\mu}]\mathbf{x} \right)_i \boldsymbol{\phi}_i, \tag{4.29b}$$

their "transposed" versions

$$\mathbb{P}_h^T[\boldsymbol{\mu}] : \boldsymbol{\mathcal{S}}_h \to \left(\mathcal{S}_h \right)^2 : \boldsymbol{v} \to \sum_{i=1}^N \left(\mathbf{H}[\boldsymbol{\mu}(z_i)]^T \mathbf{T}\boldsymbol{v}(z_i) \right) \varphi_i \tag{4.30a}$$

$$\widetilde{\mathbb{P}}_h^T[\boldsymbol{\mu}] : \mathbb{R}^{3N} \to \left(\mathcal{S}_h \right)^2 : \mathbf{x} \mapsto \sum_{i=1}^{3N} \left(\mathbf{Q}[\boldsymbol{\mu}]^T \mathbf{x} \right)_i \boldsymbol{\phi}_i, \tag{4.30b}$$

and the compositions

$$\boldsymbol{\Pi}_h[\boldsymbol{\mu}] : \boldsymbol{\mathcal{S}}_h \to \mathcal{K}_h[\boldsymbol{\mu}] \subsetneq \boldsymbol{\mathcal{S}}_h : \boldsymbol{v} \mapsto \sum_{i=1}^N \left(\mathbf{TH}[\mathbf{T}\boldsymbol{\mu}(z_i)]\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z_i)]^T \mathbf{T}\boldsymbol{v}(z_i) \right) \varphi_i, \tag{4.31a}$$

$$\widetilde{\boldsymbol{\Pi}}_h[\boldsymbol{\mu}] : \mathbb{R}^{3N} \to \mathcal{K}_h[\boldsymbol{\mu}] \subsetneq \boldsymbol{\mathcal{S}}_h : \mathbf{x} \mapsto \sum_{i=1}^{3N} \left(\mathbf{Q}[\boldsymbol{\mu}]\mathbf{Q}[\boldsymbol{\mu}]^T \mathbf{x} \right)_i \boldsymbol{\phi}_i. \tag{4.31b}$$

Note that $\boldsymbol{\Pi}_h[\boldsymbol{\mu}]$ is the nodewise orthogonal projection onto $\mathcal{K}_h[\boldsymbol{\mu}]$. The following lemma discusses the relations of the mappings (4.29)–(4.31).

Lemma 4.6.1. *For any $\boldsymbol{\mu} \in \boldsymbol{\mathcal{M}}_h$, there hold the following assertions (i)–(v):*
(i) *For $\mathbf{x} \in \mathbb{R}^{2N}$ and $\boldsymbol{w} := \sum_{i=1}^{2N} \mathbf{x}_i \boldsymbol{\psi}_i \in (\mathcal{S}_h)^2$, it holds that $\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w} = \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}$.*
(ii) *For $\mathbf{y} \in \mathbb{R}^{3N}$ and $\boldsymbol{v} := \sum_{i=1}^{3N} \mathbf{y}_i \boldsymbol{\phi}_i$, it holds that $\boldsymbol{\Pi}_h[\boldsymbol{\mu}]\boldsymbol{v} = \widetilde{\boldsymbol{\Pi}}_h[\boldsymbol{\mu}]\mathbf{y}$.*
(iii) *For $\boldsymbol{v} \in \boldsymbol{\mathcal{S}}_h$, it holds that $\mathbb{P}_h[\boldsymbol{\mu}] \circ \mathbb{P}_h^T[\boldsymbol{\mu}]\,\boldsymbol{v} = \boldsymbol{\Pi}_h[\boldsymbol{\mu}]\,\boldsymbol{v}$.*
(iv) *For $\boldsymbol{w} \in (\mathcal{S}_h)^2$, it holds that $\mathbb{P}_h^T[\boldsymbol{\mu}] \circ \mathbb{P}_h[\boldsymbol{\mu}]\,\boldsymbol{w} = \boldsymbol{w}$.*
(v) *$\mathbb{P}_h[\boldsymbol{\mu}] : (\mathcal{S}_h)^2 \to \boldsymbol{\mathcal{K}}_h[\boldsymbol{\mu}]$ and $\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}] : \mathbb{R}^{2N} \to \boldsymbol{\mathcal{K}}_h[\boldsymbol{\mu}]$ are isomorphisms.*

Proof. (i)–(ii) follow by definition. (iii)–(iv) follow from the block-structure of $\mathbf{Q}[\boldsymbol{\mu}]$, since $\mathbf{Q}[\boldsymbol{\mu}]$ has orthonormal columns. Since $(\mathcal{S}_h)^2 \cong \mathbb{R}^{2N} \cong \boldsymbol{\mathcal{K}}_h[\boldsymbol{\mu}]$, (iv) proves that $\mathbb{P}_h[\boldsymbol{\mu}]$ is an isomorphism. Together with (i), this also proves the statement about $\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]$. Altogether, this concludes the proof. $\qquad\square$

4.6.2 Proof of Theorem 4.3.1

Since $\mathbf{A}[\boldsymbol{m}^n]$ is positive definite and $\mathbf{Q}[\boldsymbol{m}^n]$ has orthonormal columns, the system matrix in (4.12) is also positive definite. Let $\mathbf{x} \in \mathbb{R}^{2N}$ be the unique solution of (4.12). Then, it holds that

$$\mathbf{Q}[\boldsymbol{m}^n]^T \mathbf{A}[\boldsymbol{m}^n] \mathbf{Q}[\boldsymbol{m}^n]\mathbf{x} \cdot \mathbf{y} \overset{(4.12)}{=} \mathbf{Q}[\boldsymbol{m}^n]^T \mathbf{b}[\boldsymbol{m}^n] \cdot \mathbf{y} \quad \text{for all } \mathbf{y} \in \mathbb{R}^{2N}. \tag{4.32}$$

We denote the bilinear form on the left-hand side of (4.5) by $A_h(\cdot, \cdot)$ and the linear functional on the right-hand side of (4.5) by $R(\cdot)$. The definition of $\mathbf{A}[\boldsymbol{m}^n]$ in Section 4.2.4 then yields that

$$A_h(\widetilde{\mathbb{P}}_h[\boldsymbol{m}^n]\mathbf{x}, \widetilde{\mathbb{P}}_h[\boldsymbol{m}^n]\mathbf{y}) = \mathbf{A}[\boldsymbol{m}^n]\mathbf{Q}[\boldsymbol{m}^n]\mathbf{x} \cdot \mathbf{Q}[\boldsymbol{m}^n]\mathbf{y} = \mathbf{Q}[\boldsymbol{m}^n]^T \mathbf{A}[\boldsymbol{m}^n]\mathbf{Q}[\boldsymbol{m}^n]\mathbf{x} \cdot \mathbf{y}$$

$$\overset{(4.32)}{=} \mathbf{Q}[\boldsymbol{m}^n]^T \mathbf{b}[\boldsymbol{m}^n] \cdot \mathbf{y} = \mathbf{b}[\boldsymbol{m}^n] \cdot \mathbf{Q}[\boldsymbol{m}^n]\mathbf{y} = R(\widetilde{\mathbb{P}}_h[\boldsymbol{m}^n]\mathbf{y}) \quad \text{for all } \mathbf{y} \in \mathbb{R}^{2N}.$$

With Lemma 4.6.1 (v), $\widetilde{\mathbb{P}}_h[\boldsymbol{m}^n]$ is an isomorphism from $\mathbb{R}^{2N}$ to $\boldsymbol{\mathcal{K}}_h[\boldsymbol{m}^n]$. Consequently, the function $\widetilde{\mathbb{P}}_h[\boldsymbol{m}^n]\mathbf{x} \in \boldsymbol{\mathcal{K}}_h[\boldsymbol{m}^n]$ is a solution to (4.5). The representation formula (4.13) follows from (4.29b). This concludes the proof. $\qquad\square$

4.6.3 Stability analysis of auxiliary mappings

The first lemma proves discrete $\boldsymbol{L}^2$-stabilities of the mappings (4.29)–(4.31).

Lemma 4.6.2. *There exists $C > 0$, which depends only on $C_{\mathbf{mesh}}$, such that the following assertions (i)–(vi) hold true:*
(i) *For any $\boldsymbol{\mu} \in \boldsymbol{\mathcal{M}}_h$, it holds that*

$$C^{-1}\|\boldsymbol{w}\|_\Omega \le \|\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w}\|_\Omega \le C\|\boldsymbol{w}\|_\Omega \quad \text{for all } \boldsymbol{w} \in (\mathcal{S}_h)^2.$$

(ii) *For any $\boldsymbol{\mu} \in \boldsymbol{\mathcal{M}}_h$, it holds that*

$$\|\mathbb{P}_h^T[\boldsymbol{\mu}]\boldsymbol{v}\|_\Omega \le C\,\|\boldsymbol{v}\|_\Omega \quad \text{for all } \boldsymbol{v} \in \boldsymbol{\mathcal{S}}_h.$$

(iii) *For any $\boldsymbol{\mu} \in \boldsymbol{\mathcal{M}}_h$, it holds that*

$$\|\boldsymbol{\Pi}_h[\boldsymbol{\mu}]\boldsymbol{v}\|_\Omega \le C\|\boldsymbol{v}\|_\Omega \quad \text{for all } \boldsymbol{v} \in \boldsymbol{\mathcal{S}}_h.$$

(iv) For any $\boldsymbol{\mu} \in \mathcal{M}_h$, it holds that

$$C^{-1}h^{3/2}|\mathbf{x}| \leq \|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\|_\Omega \leq Ch^{3/2}|\mathbf{x}| \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

(v) For any $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{M}_h$, it holds that

$$C^{-1}\|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega \leq \|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\|_\Omega \leq C\|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

(vi) For any $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{M}_h$, it holds that

$$\|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x} - \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega \leq Ch^{3/2} \max_{z \in \mathcal{N}_h} |\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z)] - \mathbf{H}[\mathbf{T}\boldsymbol{\nu}(z)]| \, |\mathbf{x}| \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

Proof. Throughout the proof, recall that $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$. For the proof of (i), and (iv)–(vi), let $\mathbf{x} \in \mathbb{R}^{2N}$ and define $\boldsymbol{w} := \sum_{i=1}^{2N} \mathbf{x}_i \boldsymbol{\psi}_i$. For $d = 2, 3$ a scaling argument (see, e.g., [Bar15, Lemma 3.9]) yields that

$$\|\boldsymbol{\varphi}\|_\Omega^2 \leq h^3 \sum_{i=1}^N |\boldsymbol{\varphi}(z_i)|^2 \leq (d+2)\|\boldsymbol{\varphi}\|_\Omega^2 \quad \text{for all } \boldsymbol{\varphi} \in (\mathcal{S}_h)^d. \tag{4.33}$$

Since the matrices $\mathbf{H}[\boldsymbol{\mu}(z_i)] \in \mathbb{R}^{3\times2}$ have orthonormal columns, Lemma 4.6.1 (i) yields that

$$\|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\|_\Omega^2 = \|\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w}\|_\Omega^2 \stackrel{(4.33)}{\simeq} h^3 \sum_{i=1}^N |\mathbf{T}\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z_i)]\boldsymbol{w}(z_i)|^2 = h^3 \sum_{i=1}^N |\boldsymbol{w}(z_i)|^2 \stackrel{(4.33)}{\simeq} \|\boldsymbol{w}\|_\Omega^2.$$

This proves (i), and (iv) follows from $\sum_{i=1}^N |\boldsymbol{w}(z_i)|^2 = \sum_{i=1}^{2N} |\mathbf{x}_i|^2$. (v) is a direct consequence of (iv). For the proof of (vi), note that

$$\|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x} - \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega^2 \stackrel{(4.33)}{\simeq} h^3 \sum_{i=1}^N \Big| \mathbf{T}\big(\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z_i)] - \mathbf{H}[\mathbf{T}\boldsymbol{\nu}(z_i)]\big)\boldsymbol{w}(z_i) \Big|^2$$

$$\leq h^3 \max_{z \in \mathcal{N}_h} |\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z)] - \mathbf{H}[\mathbf{T}\boldsymbol{\nu}(z)]|^2 \, |\mathbf{x}|^2.$$

This proves (vi). For the proof of (ii)–(iii), let $\boldsymbol{v} \in \mathcal{S}_h$. Since the matrices $\mathbf{H}[\boldsymbol{\mu}(z_i)] \in \mathbb{R}^{3\times2}$ have orthonormal columns, we obtain that

$$\|\mathbb{P}_h^T[\boldsymbol{\mu}]\boldsymbol{v}\|_\Omega^2 \stackrel{(4.33)}{\simeq} h^3 \sum_{i=1}^N |\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z_i)]^T\mathbf{T}\boldsymbol{v}(z_i)|^2 \leq h^3 \sum_{i=1}^N |\boldsymbol{v}(z_i)|^2 \stackrel{(4.33)}{\simeq} \|\boldsymbol{v}\|_\Omega^2.$$

This proves (ii). Together with (i)–(ii) and Lemma 4.6.1(iii), this also proves (iii). Altogether, this concludes the proof. $\square$

The second lemma proves discrete $\boldsymbol{H}^1$-stability properties of the mappings (4.29)–(4.31). In contrast to Lemma 4.6.1 and Lemma 4.6.2, the proof exploits the explicit definition of the Householder matrices (4.10).

Lemma 4.6.3. *Let $\boldsymbol{\mu} \in \mathcal{M}_h$ with $1 + (\boldsymbol{\mu}(z))_3 \geq \gamma > 0$ for all $z \in \mathcal{N}_h$. Then, there exists $C > 1$, which depends only on $C_{\mathbf{mesh}}$, such that the following assertions (i)–(iii) hold true:*
(i) For all $\boldsymbol{w} \in (\mathcal{S}_h)^2$, it holds that

$$\|\nabla\big(\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w}\big)\|_\Omega \leq C\gamma^{-2}\,\|\nabla\boldsymbol{\mu}\|_\infty\,\|\boldsymbol{w}\|_\Omega + C\,\|\nabla\boldsymbol{w}\|_\Omega.$$

(ii) For all $\boldsymbol{v} \in \mathcal{S}_h$, it holds that

$$\|\nabla\big(\mathbb{P}_h^T[\boldsymbol{\mu}]\boldsymbol{v}\big)\|_\Omega \leq C\gamma^{-2}\,\|\nabla\boldsymbol{\mu}\|_\infty\,\|\boldsymbol{v}\|_\Omega + C\,\|\nabla\boldsymbol{v}\|_\Omega.$$

(iii) For all $\boldsymbol{v} \in \mathcal{S}_h$, it holds that

$$\|\nabla\big(\boldsymbol{\Pi}_h[\boldsymbol{\mu}]\boldsymbol{v}\big)\|_\Omega \leq C\gamma^{-2}\,\|\nabla\boldsymbol{\mu}\|_\infty\,\|\boldsymbol{v}\|_\Omega + C\,\|\nabla\boldsymbol{v}\|_\Omega.$$

Proof. First, we prove (i). We split the proof into the following six steps.

Step 1. We derive a handier representation of $\mathbb{P}_h[\boldsymbol{\mu}]$. Let $\mu_1, \mu_2, \mu_3 \in \mathcal{S}_h$ such that $\mathbf{T}\boldsymbol{\mu} := (\mu_1, \mu_2, \mu_3)^T$. Since functions in $\mathcal{S}_h$ attain their minimum in one of the nodes, we obtain, in particular, that

$$1 + \mu_3 = 1 + (\mathbf{T}\boldsymbol{\mu})_3 \geq \gamma > 0 \quad \text{in } \Omega. \tag{4.34}$$

Hence, we can interpret

$$\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\mu_1 & -\mu_2 \end{pmatrix}, \quad \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] := \frac{1}{1+\mu_3}\begin{pmatrix} \mu_1^2 & \mu_1\mu_2 \\ \mu_1\mu_2 & \mu_2^2 \\ 0 & 0 \end{pmatrix} \tag{4.35}$$

and

$$\boldsymbol{q}[\boldsymbol{\mu}] := \mathbf{T}\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T}\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] \tag{4.36}$$

as functions $\boldsymbol{q}[\boldsymbol{\mu}], \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}], \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] : \Omega \to \mathbb{R}^{3\times2}$. With the definition of the Householder matrices (4.10), an elementary calculation shows that $\mathbf{T}\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z_i)] = \boldsymbol{q}[\boldsymbol{\mu}(z_i)]$ for all $i = 1, \ldots, N$. With $\mathcal{I}_h$ being the vector-valued nodal interpolant onto $\mathcal{S}_h$, we get

$$\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w} \overset{(4.29)}{=} \mathcal{I}_h(\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w}) \quad \text{for all } \boldsymbol{w} \in (\mathcal{S}_h)^2. \tag{4.37}$$

Step 2. We derive preliminary estimates for $\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}]$ and $\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}]$ from (4.35). To this end, recall that $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$. Lemma 4.6.9(i) yields that

$$\|\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}]\|_\infty \lesssim 1 \quad \text{and} \quad \|\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}]\|_\infty \lesssim 1. \tag{4.38a}$$

Let $k \in \{1, 2, 3\}$. Lemma 4.6.9(ii) yields that

$$\|\partial_k\big[\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}]\big]\|_\infty \overset{(4.34)}{\lesssim} \|\nabla\mathbf{T}\boldsymbol{\mu}\|_\infty = \|\mathbf{T}\nabla\boldsymbol{\mu}\|_\infty = \|\nabla\boldsymbol{\mu}\|_\infty, \tag{4.38b}$$

as well as

$$\|\partial_k\big[\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}]\big]\|_\infty \overset{(4.34)}{\lesssim} \gamma^{-1}\|\nabla\mathbf{T}\boldsymbol{\mu}\|_\infty = \gamma^{-1}\|\nabla\boldsymbol{\mu}\|_\infty. \tag{4.38c}$$

Let $\ell, k \in \{1, 2, 3\}$. The definition (4.35) and Lemma 4.6.9(iii) yield that, elementwise,

$$\partial_\ell \partial_k \big[\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \big] = \mathbf{0} \tag{4.38d}$$

as well as

$$\max_{K \in \mathcal{T}_h} \| \partial_\ell \partial_k \big[\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \|_{\boldsymbol{L}^\infty(K)} \overset{(4.34)}{\lesssim} \gamma^{-2} \| \nabla \mathbf{T}\boldsymbol{\mu} \|_\infty^2 = \gamma^{-2} \| \nabla \boldsymbol{\mu} \|_\infty^2. \tag{4.38e}$$

Step 3. For $\boldsymbol{w} \in (\mathcal{S}_h)^2$, we estimate $\| \nabla [\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w}] \|_\Omega$. To that end, note that $\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w}|_K \in (H^2(K))^3$ for all elements K. We exploit the elementwise approximation properties of the nodal interpolant $\boldsymbol{\mathcal{I}}_h$ and obtain that

$$\| \nabla \big[\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{v} \big] \|_\Omega \overset{(4.29)}{=} \| \nabla \big[\boldsymbol{\mathcal{I}}_h (\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w}) \big] \|_\Omega \lesssim \| \nabla \big[\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w} \big] \|_\Omega + \| \nabla \big[(\mathbf{1} - \boldsymbol{\mathcal{I}}_h)(\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w}) \big] \|_\Omega$$

$$\lesssim \| \nabla \big[\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w} \big] \|_\Omega + h \Big(\sum_{K \in \mathcal{T}_h} \| D^2 \big[(\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w}) \big] \|_K^2 \Big)^{1/2} =: T_1 + h \, T_2. \tag{4.39}$$

Step 4. We estimate T_1. Let $k \in \{1, 2, 3\}$. The product rule yields that

$$\partial_k \big[\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w} \big] = \partial_k \big[\boldsymbol{q}[\boldsymbol{\mu}] \big] \boldsymbol{w} + \boldsymbol{q}[\boldsymbol{\mu}] \partial_k \boldsymbol{w}$$

$$\overset{(4.36)}{=} \partial_k \big[\mathbf{T}\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \big] \boldsymbol{w} - \partial_k \big[\mathbf{T}\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \boldsymbol{w} + \mathbf{T}\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \partial_k \boldsymbol{w} - \mathbf{T}\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \partial_k \boldsymbol{w}$$

$$= \mathbf{T} \partial_k \big[\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \big] \boldsymbol{w} - \mathbf{T} \partial_k \big[\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \boldsymbol{w} + \mathbf{T}\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \partial_k \boldsymbol{w} - \mathbf{T}\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \partial_k \boldsymbol{w}. \tag{4.40}$$

Note that $\gamma \leq 1 + \mu_3 \leq 1 + |\mu_3| \leq 2$ and recall that $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$. With the estimates from (4.38) and with $1 \leq 2/\gamma$, the latter equation yields that

$$T_1 \overset{(4.39)}{\lesssim} \gamma^{-1} \| \nabla \mathbf{T}\boldsymbol{\mu} \|_\infty \| \boldsymbol{w} \|_\Omega + \| \nabla \boldsymbol{w} \|_\Omega = \gamma^{-1} \| \nabla \boldsymbol{\mu} \|_\infty \| \boldsymbol{w} \|_\Omega + \| \nabla \boldsymbol{w} \|_\Omega.$$

Step 5. We estimate T_2. Let $\ell, k \in \{1, 2, 3\}$. Elementwise, it holds that

$$\partial_\ell \mathbf{T} \partial_k \mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] = \mathbf{T} \partial_\ell \partial_k \mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \overset{(4.38d)}{=} \mathbf{0}$$

as well as $\partial_\ell \partial_k \boldsymbol{w} = \mathbf{0}$. Together with the product rule, this yields that

$$\partial_\ell \partial_k \big[\boldsymbol{q}[\boldsymbol{\mu}]\boldsymbol{w} \big] \overset{(4.40)}{=} \partial_k \big[\mathbf{T}\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_\ell \boldsymbol{w} - \partial_\ell \mathbf{T} \partial_k \big[\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \boldsymbol{w} - \partial_k \big[\mathbf{T}\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_\ell \boldsymbol{w}$$

$$+ \partial_\ell \mathbf{T} \big[\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_k \boldsymbol{w} - \partial_\ell \big[\mathbf{T}\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_k \boldsymbol{w}$$

$$= \mathbf{T} \partial_k \big[\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_\ell \boldsymbol{w} - \mathbf{T} \partial_\ell \partial_k \big[\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \boldsymbol{w} - \mathbf{T} \partial_k \big[\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_\ell \boldsymbol{w}$$

$$+ \mathbf{T} \partial_\ell \big[\mathbf{R}_1 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_k \boldsymbol{w} - \mathbf{T} \partial_\ell \big[\mathbf{R}_2 [\mathbf{T}\boldsymbol{\mu}] \big] \partial_k \boldsymbol{w}.$$

Recall that $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$. With (4.38) and $1 \leq 2/\gamma$, the latter equation yields that

$$T_2 \overset{(4.39)}{\lesssim} \gamma^{-1} \| \nabla \mathbf{T}\boldsymbol{\mu} \|_\infty \| \nabla \boldsymbol{w} \|_\Omega + \gamma^{-2} \| \nabla \mathbf{T}\boldsymbol{\mu} \|_\infty^2 \| \boldsymbol{w} \|_\Omega$$

$$= \gamma^{-1} \| \nabla \boldsymbol{\mu} \|_\infty \| \nabla \boldsymbol{w} \|_\Omega + \gamma^{-2} \| \nabla \boldsymbol{\mu} \|_\infty^2 \| \boldsymbol{w} \|_\Omega.$$

Step 6. We combine **Step 3**–**Step 5**. For all $\boldsymbol{w} \in (\mathcal{S}_h)^2$, this yields that

$$\|\nabla[\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w}]\|_\Omega \lesssim \gamma^{-1} \|\nabla\boldsymbol{\mu}\|_\infty \|\boldsymbol{w}\|_\Omega + \|\nabla\boldsymbol{w}\|_\Omega$$
$$+ h\,\gamma^{-1} \|\nabla\boldsymbol{\mu}\|_\infty \|\nabla\boldsymbol{w}\|_\Omega + h\,\gamma^{-2} \|\nabla\boldsymbol{\mu}\|_\infty^2 \|\boldsymbol{w}\|_\Omega.$$

With an inverse estimate and $1 \le 2/\gamma$, the latter equation yields for all $\boldsymbol{w} \in (\mathcal{S}_h)^2$ that

$$\|\nabla[\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w}]\|_\Omega \lesssim \gamma^{-2} \|\nabla\boldsymbol{\mu}\|_\infty \|\boldsymbol{w}\|_\Omega + \|\nabla\boldsymbol{w}\|_\Omega. \tag{4.41}$$

This concludes the proof of (i).

For the proof of (ii), let $\widetilde{\boldsymbol{\mathcal{I}}}_h$ be the nodal interpolant in 2D. With $\widetilde{\boldsymbol{\mathcal{I}}}_h$ instead of $\boldsymbol{\mathcal{I}}_h$ and $\boldsymbol{q}[\boldsymbol{\mu}]^T : \Omega \to \mathbb{R}^{2\times3}$ instead of $\boldsymbol{q}[\boldsymbol{\mu}]$, the proof of (ii) follows the lines of **Step 1**–**Step 5**.

For the proof of (iii), let $\boldsymbol{v} \in \mathcal{S}_h$ and $\boldsymbol{w} := \mathbb{P}_h^T[\boldsymbol{\mu}]\boldsymbol{v} \in \mathcal{S}_h$. With Lemma 4.6.1(iii) and Lemma 4.6.2(ii), we get that

$$\|\nabla[\boldsymbol{\Pi}_h[\boldsymbol{\mu}]\boldsymbol{v}]\|_\Omega = \|\nabla\left[\left(\mathbb{P}_h[\boldsymbol{\mu}] \circ \mathbb{P}_h^T[\boldsymbol{\mu}]\right)\boldsymbol{v}\right]\|_\Omega \lesssim \|\nabla\left[\mathbb{P}_h[\boldsymbol{\mu}]\boldsymbol{w}\right]\|_\Omega$$

$$\overset{\text{(i)}}{\lesssim} \gamma^{-2} \|\nabla\boldsymbol{\mu}\|_\infty \|\mathbb{P}_h^T[\boldsymbol{\mu}]\boldsymbol{v}\|_\Omega + \|\nabla[\mathbb{P}_h^T[\boldsymbol{\mu}]\boldsymbol{v}]\|_\Omega$$

$$\lesssim \gamma^{-2}\|\nabla\boldsymbol{\mu}\|_\infty \|\boldsymbol{v}\|_\Omega + \|\nabla[\mathbb{P}_h^T[\boldsymbol{\mu}]\boldsymbol{v}]\|_\Omega$$

$$\overset{\text{(ii)}}{\lesssim} \gamma^{-2} \|\nabla\boldsymbol{\mu}\|_\infty \|\boldsymbol{v}\|_\Omega + \|\nabla\boldsymbol{v}\|_\Omega.$$

This proves (iii) and concludes the proof. $\qquad\square$

In the final auxiliary lemma, we prove a discrete $\boldsymbol{H}^1$-continuity of the mapping $\mathbb{P}_h(\cdot)$ from (4.29). Unlike Lemma 4.6.1 and Lemma 4.6.2, the following lemma builds on the explicit definition of the Householder matrices (4.10).

Lemma 4.6.4. *Let* $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{M}_h$ *with* $1 + (\mathbf{T}\boldsymbol{\mu}(z))_3 \ge \gamma > 0$ *and* $1 + (\mathbf{T}\boldsymbol{\nu}(z))_3 \ge \gamma > 0$ *for all* $z \in \mathcal{N}_h$. *Then, there exists* $C > 1$, *which depends only on* $C_{\mathbf{mesh}}$, *such that*

$$\|\nabla\left(\left(\mathbb{P}_h[\boldsymbol{\mu}]-\mathbb{P}_h[\boldsymbol{\nu}]\right)\boldsymbol{w}\right)\|_\Omega \le C\,\gamma^{-1} \|\nabla\boldsymbol{\mu}-\nabla\boldsymbol{\nu}\|_\infty \|\boldsymbol{w}\|_\Omega + C\,\gamma^{-1} \|\boldsymbol{\mu}-\boldsymbol{\nu}\|_\infty \|\nabla\boldsymbol{w}\|_\Omega$$
$$+ C\,\gamma^{-3}\left(\|\nabla\boldsymbol{\mu}\|_\infty+\|\nabla\boldsymbol{\nu}\|_\infty\right)\|\boldsymbol{\mu}-\boldsymbol{\nu}\|_\infty \|\boldsymbol{w}\|_\Omega, \tag{4.42}$$

for all $\boldsymbol{w} \in \left(\mathcal{S}_h\right)^2$.

Proof. We split the proof into the following six steps.

Step 1. With the assumption $1 + (\mathbf{T}\boldsymbol{\mu}(z))_3 \ge \gamma > 0$ and $1 + (\mathbf{T}\boldsymbol{\nu}(z))_3 \ge \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, we use the definitions (4.35) of $\mathbf{R}_1(\cdot)$ and $\mathbf{R}_2(\cdot)$ and interpret

$$\boldsymbol{q}[\boldsymbol{\mu}] := \mathbf{T}\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T}\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}], \tag{4.43a}$$
$$\boldsymbol{q}[\boldsymbol{\nu}] := \mathbf{T}\mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}] - \mathbf{T}\mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}] \tag{4.43b}$$

as functions $\boldsymbol{q}[\boldsymbol{\mu}], \boldsymbol{q}[\boldsymbol{\nu}] : \Omega \to \mathbb{R}^{3\times2}$. With $\boldsymbol{\mathcal{I}}_h$ being the vector-valued nodal interpolant onto $\mathcal{S}_h$, recall from (4.37) that

$$\left(\mathbb{P}_h[\boldsymbol{\mu}] - \mathbb{P}_h[\boldsymbol{\nu}]\right)\boldsymbol{w} = \boldsymbol{\mathcal{I}}_h\left(\left(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}]\right)\boldsymbol{w}\right) \quad \text{for all } \boldsymbol{w} \in (\mathcal{S}_h)^2. \tag{4.44}$$

Step 2. Recall that $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$. With the definition (4.35) of $\mathbf{R}_1(\cdot)$, we get that

$$\|\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}]\|_\infty \lesssim \|\mathbf{T}\boldsymbol{\mu} - \mathbf{T}\boldsymbol{\nu}\|_\infty = \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty. \tag{4.45a}$$

With the product rule, we further get for all $\ell, k \in \{1, 2, 3\}$ that

$$\begin{aligned}
\|-\|_{\partial_k \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}} \partial_k \mathbf{R}_1 \infty][\mathbf{T}\boldsymbol{\nu}] &\lesssim \|\nabla \mathbf{T}\boldsymbol{\mu} - \nabla \mathbf{T}\boldsymbol{\nu}\|_\infty \\
&= \|\mathbf{T}\nabla\boldsymbol{\mu} - \mathbf{T}\nabla\boldsymbol{\nu}\|_\infty = \|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty
\end{aligned} \tag{4.45b}$$

as well as

$$\partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] = \mathbf{0} = \partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}]. \tag{4.45c}$$

Moreover, define

$$\begin{aligned}
\sigma(\mathbf{T}\boldsymbol{\mu}, \mathbf{T}\boldsymbol{\nu}) := \|\nabla \mathbf{T}\boldsymbol{\mu}\|_\infty + \|\nabla \mathbf{T}\boldsymbol{\nu}\|_\infty &= \|\mathbf{T}\nabla\boldsymbol{\mu}\|_\infty + \|\mathbf{T}\nabla\boldsymbol{\nu}\|_\infty \\
&= \|\nabla\boldsymbol{\mu}\|_\infty + \|\nabla\boldsymbol{\nu}\|_\infty = \sigma(\boldsymbol{\mu}, \boldsymbol{\nu}).
\end{aligned} \tag{4.46a}$$

Note that an inverse inequality yields that

$$h\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu}) \lesssim \|\boldsymbol{\mu}\|_\infty + \|\boldsymbol{\nu}\|_\infty = 2. \tag{4.46b}$$

Lemma 4.6.10 and the definition (4.35) of $\mathbf{R}_2(\cdot)$ then yield that

$$\|\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}]\|_\infty \lesssim \gamma^{-1}\|\mathbf{T}\boldsymbol{\mu} - \mathbf{T}\boldsymbol{\nu}\|_\infty = \gamma^{-1}\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty. \tag{4.47a}$$

For all $\ell, k \in \{1, 2, 3\}$, we further get that

$$\begin{aligned}
&\|\partial_k \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \partial_k \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}]\|_\infty \\
&\lesssim \gamma^{-2}\,\sigma(\mathbf{T}\boldsymbol{\mu}, \mathbf{T}\boldsymbol{\nu})\,\|\mathbf{T}\boldsymbol{\mu} - \mathbf{T}\boldsymbol{\nu}\|_\infty + \gamma^{-1}\|\nabla \mathbf{T}\boldsymbol{\mu} - \nabla \mathbf{T}\boldsymbol{\nu}\|_\infty \\
&\overset{(4.46)}{=} \gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\mathbf{T}\boldsymbol{\mu} - \mathbf{T}\boldsymbol{\nu}\|_\infty + \gamma^{-1}\|\mathbf{T}\nabla\boldsymbol{\mu} - \mathbf{T}\nabla\boldsymbol{\nu}\|_\infty \\
&= \gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty + \gamma^{-1}\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty
\end{aligned} \tag{4.47b}$$

as well as

$$\begin{aligned}
&\|\partial_\ell \partial_k \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \partial_\ell \partial_k \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}]\|_\infty \\
&\lesssim \gamma^{-3}\,\sigma(\mathbf{T}\boldsymbol{\mu}, \mathbf{T}\boldsymbol{\nu})^2\,\|\mathbf{T}\boldsymbol{\mu} - \mathbf{T}\boldsymbol{\nu}\|_\infty + \gamma^{-2}\,\sigma(\mathbf{T}\boldsymbol{\mu}, \mathbf{T}\boldsymbol{\nu})\,\|\nabla \mathbf{T}\boldsymbol{\mu} - \nabla \mathbf{T}\boldsymbol{\nu}\|_\infty \\
&\overset{(4.46)}{=} \gamma^{-3}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})^2\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty + \gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty.
\end{aligned} \tag{4.47c}$$

Step 3. Let $\boldsymbol{w} \in (\mathcal{S}_h)^2$. Standard estimates for the nodal interpolant $\mathcal{I}_h$ yield that

$$\begin{aligned}
\|\nabla\big[(\mathbb{P}_h[\boldsymbol{\mu}] - \mathbb{P}_h[\boldsymbol{\nu}])\,\boldsymbol{w}\big]\|_\Omega &\overset{(4.44)}{=} \|\nabla\mathcal{I}_h\big[(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}])\,\boldsymbol{w}\big]\|_\Omega \\
&\leq \|\nabla\big[(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}])\,\boldsymbol{w}\big]\|_\Omega + \|\nabla(\mathbf{1} - \mathcal{I}_h)\big[(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}])\,\boldsymbol{w}\big]\|_\Omega \\
&\lesssim \|\nabla\big[(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}])\,\boldsymbol{w}\big]\|_\Omega + h\left(\sum_{K \in \mathcal{T}_h} \|D^2\big[(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}])\,\boldsymbol{w}\big]\|_K^2\right)^{1/2} \\
&=: T_1 + h\,T_2.
\end{aligned}$$

Step 4. We estimate T_1. Let $k \in \{1, 2, 3\}$. With the product rule, we get that

$$
\begin{aligned}
\partial_k \big((\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}]) \, \boldsymbol{w} \big) \;=\;& \partial_k \big((\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}]) \, \big) \boldsymbol{w} + (\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}]) \, \partial_k \boldsymbol{w} \\
\overset{(4.43)}{=}\;& \partial_k \big((\mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}]) \, \big) \boldsymbol{w} - \partial_k \big((\mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}]) \, \big) \boldsymbol{w} \\
& + (\mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}]) \, \partial_k \boldsymbol{w} - (\mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}]) \, \partial_k \boldsymbol{w} \\
\;=\;& \mathbf{T} \partial_k \big((\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}]) \, \big) \boldsymbol{w} - \mathbf{T} \partial_k \big((\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}]) \, \big) \boldsymbol{w} \\
& + \mathbf{T} (\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}]) \, \partial_k \boldsymbol{w} - \mathbf{T} (\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}]) \, \partial_k \boldsymbol{w}.
\end{aligned}
$$

Recall that $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$. With the estimates from **Step 2**, we further get that

$$
\begin{aligned}
\| \partial_k \big((\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}]) \, \boldsymbol{w} \big) \|_\Omega \;\leq\;& \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \boldsymbol{w} \|_\Omega + \gamma^{-2} \, \sigma(\boldsymbol{\mu}, \boldsymbol{\nu}) \, \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \boldsymbol{w} \|_\Omega \\
& + \gamma^{-1} \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \boldsymbol{w} \|_\Omega + \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \partial_k \boldsymbol{w} \|_\Omega \\
& + \gamma^{-1} \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \partial_k \boldsymbol{w} \|_\Omega.
\end{aligned}
$$

With $1 \leq 2/\gamma$, we arrive at

$$
\begin{aligned}
T_1 \;\lesssim\;& \gamma^{-1} \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \boldsymbol{w} \|_\Omega + \gamma^{-1} \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \nabla \boldsymbol{w} \|_\Omega \\
& + \gamma^{-2} \, \sigma(\boldsymbol{\mu}, \boldsymbol{\nu}) \, \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \boldsymbol{w} \|_\Omega.
\end{aligned}
$$

Step 5. We estimate T_2. To this end, let $\ell, k \in \{1, 2, 3\}$. Note that the second derivative of the piecewise affine function $\boldsymbol{w}$ vanishes on each element $K \in \mathcal{T}_h$. Moreover, recall from (4.45c) that elementwise $\partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] = \mathbf{0} = \partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}]$. The product rule yields elementwise that

$$
\begin{aligned}
& \partial_\ell \partial_k \big((\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}]) \, \boldsymbol{w} \big) \\
& =\; \partial_\ell \partial_k \big(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}] \big) \boldsymbol{w} + \partial_k \big(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}] \big) \partial_\ell \boldsymbol{w} + \partial_\ell \big(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}] \big) \partial_k \boldsymbol{w} \\
& \overset{(4.43)}{=}\; \partial_\ell \partial_k \big(\mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}] \big) \boldsymbol{w} + \partial_k \big(\mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}] \big) \partial_\ell \boldsymbol{w} \\
& \qquad - \partial_k \big(\mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}] \big) \partial_\ell \boldsymbol{w} + \partial_\ell \big(\mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}] \big) \partial_k \boldsymbol{w} \\
& \qquad - \partial_\ell \big(\mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{T} \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}] \big) \partial_k \boldsymbol{w} \\
& =\; \mathbf{T} \partial_\ell \partial_k \big(\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}] \big) \boldsymbol{w} + \mathbf{T} \partial_k \big(\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}] \big) \partial_\ell \boldsymbol{w} \\
& \qquad - \mathbf{T} \partial_k \big(\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}] \big) \partial_\ell \boldsymbol{w} + \mathbf{T} \partial_\ell \big(\mathbf{R}_1[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_1[\mathbf{T}\boldsymbol{\nu}] \big) \partial_k \boldsymbol{w} \\
& \qquad - \mathbf{T} \partial_\ell \big(\mathbf{R}_2[\mathbf{T}\boldsymbol{\mu}] - \mathbf{R}_2[\mathbf{T}\boldsymbol{\nu}] \big) \partial_k \boldsymbol{w}.
\end{aligned}
$$

With the estimates from **Step 2** and since $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$, we get for all $K \in \mathcal{T}_h$ that

$$
\begin{aligned}
& \| \partial_\ell \partial_k \big[(\boldsymbol{q}[\boldsymbol{\mu}] - \boldsymbol{q}[\boldsymbol{\nu}]) \, \boldsymbol{w} \big] \|_K \\
& \lesssim \gamma^{-3} \, \sigma(\boldsymbol{\mu}, \boldsymbol{\nu})^2 \, \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \boldsymbol{w} \|_K + \gamma^{-2} \, \sigma(\boldsymbol{\mu}, \boldsymbol{\nu}) \, \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \boldsymbol{w} \|_K \\
& \qquad + \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \partial_\ell \boldsymbol{w} \|_K + \gamma^{-1} \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \partial_\ell \boldsymbol{w} \|_K \\
& \qquad + \gamma^{-2} \, \sigma(\boldsymbol{\mu}, \boldsymbol{\nu}) \, \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \partial_\ell \boldsymbol{w} \|_K + \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \partial_k \boldsymbol{w} \|_K \\
& \qquad + \gamma^{-1} \| \nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu} \|_\infty \| \partial_k \boldsymbol{w} \|_K + \gamma^{-2} \, \sigma(\boldsymbol{\mu}, \boldsymbol{\nu}) \, \| \boldsymbol{\mu} - \boldsymbol{\nu} \|_\infty \| \partial_k \boldsymbol{w} \|_K.
\end{aligned}
$$

With $1 \leq 2/\gamma$, the latter estimate simplifies to

$$T_2 \lesssim \gamma^{-3}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})^2\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty \|\boldsymbol{w}\|_\Omega + \gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty\,\|\nabla\boldsymbol{w}\|_\Omega$$
$$+ \gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty\|\boldsymbol{w}\|_\Omega + \gamma^{-1}\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty\,\|\nabla\boldsymbol{w}\|_\Omega.$$

Step 6. We combine **Step 3**–**Step 5**. An inverse estimate and $1 \leq 2/\gamma$ imply that

$$\|\nabla\big[\,(\mathbb{P}_h[\boldsymbol{\mu}] - \mathbb{P}_h[\boldsymbol{\nu}])\,\boldsymbol{w}\,\big]\|_\Omega \leq T_1 + hT_2$$
$$\lesssim\ \gamma^{-1}\,\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty\,\|\boldsymbol{w}\|_\Omega + \gamma^{-1}\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty\,\|\nabla\boldsymbol{w}\|_\Omega$$
$$+ \gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty\,\|\boldsymbol{w}\|_\Omega + h\gamma^{-3}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})^2\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty\|\boldsymbol{w}\|_\Omega$$
$$+ h\gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty\,\|\nabla\boldsymbol{w}\|_\Omega + h\gamma^{-2}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty\|\boldsymbol{w}\|_\Omega$$
$$+ h\gamma^{-1}\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty\,\|\nabla\boldsymbol{w}\|_\Omega$$
$$\overset{(4.46)}{\lesssim}\ \gamma^{-1}\,\|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty\,\|\boldsymbol{w}\|_\Omega + \gamma^{-1}\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty\,\|\nabla\boldsymbol{w}\|_\Omega$$
$$+ \gamma^{-3}\,\sigma(\boldsymbol{\mu}, \boldsymbol{\nu})\,\|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty\|\boldsymbol{w}\|_\Omega.$$

This concludes the proof. $\qquad\square$

4.6.4 Energy norms

For $\boldsymbol{\mu} \in \mathcal{M}_h$, we define the auxiliary *theoretical* preconditioner

$$\mathbf{P}_{\mathbf{Q}}^\star[\boldsymbol{\mu}] := \big(\,\mathbf{Q}[\boldsymbol{\mu}]^T(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\text{ex}}^2 k\,\mathbf{L})\mathbf{Q}[\boldsymbol{\mu}]\,\big)^{-1} \in \mathbb{R}^{2N \times 2N}. \qquad (4.48)$$

We define the energy scalar product

$$\langle\!\langle\!\langle \mathbf{x}\,,\mathbf{y}\rangle\!\rangle\!\rangle_{\boldsymbol{\mu}}^\star := \mathbf{x} \cdot \big(\,\mathbf{P}_{\mathbf{Q}}^\star[\boldsymbol{\mu}]\,\big)^{-1}\mathbf{y} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N} \qquad (4.49)$$

and denote the induced norm by $\|\!\|\cdot\|\!\|_{\boldsymbol{\mu}}^\star$. With the definitions of the matrices from Section 4.2.4 as well as the definition (4.29b) of $\widetilde{\mathbb{P}}_h(\cdot)$, it follows that

$$\langle\!\langle\!\langle \mathbf{x}\,,\mathbf{y}\rangle\!\rangle\!\rangle_{\boldsymbol{\mu}}^\star = \alpha_{\mathbf{P}}\,\langle\Omega, \widetilde{\mathbb{P}}[\boldsymbol{\mu}]\mathbf{x}\rangle\widetilde{\mathbb{P}}[\boldsymbol{\mu}]\mathbf{y} + \ell_{\text{ex}}^2 k\,\langle\Omega, \nabla\widetilde{\mathbb{P}}[\boldsymbol{\mu}]\mathbf{x}\rangle\nabla\widetilde{\mathbb{P}}[\boldsymbol{\mu}]\mathbf{y}. \qquad (4.50)$$

This section collects equivalence results for varying arguments $\boldsymbol{\mu}$ in $\langle\!\langle\!\langle\cdot\,,\,\cdot\rangle\!\rangle\!\rangle_{\boldsymbol{\mu}}^\star$ and $\|\!\|\cdot\|\!\|_{\boldsymbol{\mu}}^\star$.

Lemma 4.6.5. *Let* $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{M}_h$ *and*

$$\widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2}) := \left(1 + \frac{\ell_{\text{ex}}^2 k}{\alpha_{\mathbf{P}} h^2}\max_{z \in \mathcal{N}_h}|\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z)] - \mathbf{H}[\mathbf{T}\boldsymbol{\nu}(z)]|^2\right)^{1/2} \geq 1. \qquad (4.51a)$$

Then, $\widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2}) = \widetilde{\kappa}(\boldsymbol{\nu}, \boldsymbol{\mu}, h^{-2})$ *and there exists a constant* $C \geq 1$*, which depends only on* $C_{\mathbf{mesh}}$*, such that*

$$C^{-1}\,\widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})^{-1}\,\|\!\|\mathbf{x}\|\!\|_{\boldsymbol{\nu}}^\star \leq \|\!\|\mathbf{x}\|\!\|_{\boldsymbol{\mu}}^\star \leq C\,\widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})\,\|\!\|\mathbf{x}\|\!\|_{\boldsymbol{\nu}}^\star \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}. \qquad (4.51b)$$

Proof. Let $\mathbf{x} \in \mathbb{R}^{2N}$. Since the symmetry $\widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2}) = \widetilde{\kappa}(\boldsymbol{\nu}, \boldsymbol{\mu}, h^{-2})$ is obvious, we only have to show that $\|\mathbf{x}\|_{\boldsymbol{\mu}}^{\star} \lesssim \widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})\|\mathbf{x}\|_{\boldsymbol{\nu}}^{\star}$. To this end, Lemma 4.6.2(v) and an inverse estimate yield that

$$(\|\mathbf{x}\|_{\boldsymbol{\mu}}^{\star})^2 \stackrel{(4.50)}{=} \alpha_{\mathbf{P}} \|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\|_{\Omega}^2 + \ell_{\mathrm{ex}}^2 k \, \|\nabla \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\|_{\Omega}^2$$

$$\lesssim \alpha_{\mathbf{P}} \|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_{\Omega}^2 + \ell_{\mathrm{ex}}^2 k \, \|\nabla \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_{\Omega}^2 + \ell_{\mathrm{ex}}^2 k \, \|\nabla \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x} - \nabla \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_{\Omega}^2$$

$$\lesssim (\|\mathbf{x}\|_{\boldsymbol{\nu}}^{\star})^2 + \ell_{\mathrm{ex}}^2 k h^{-2} \|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x} - \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_{\Omega}^2.$$

With Lemma 4.6.2(iv) and (vi), we estimate the last term by

$$\ell_{\mathrm{ex}}^2 k h^{-2} \, \|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x} - \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_{\Omega}^2 \lesssim \ell_{\mathrm{ex}}^2 k h \, |\mathbf{x}|^2 \max_{z \in \mathcal{N}_h} |\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z)] - \mathbf{H}[\mathbf{T}\boldsymbol{\nu}(z)]|^2$$

$$\simeq \ell_{\mathrm{ex}}^2 k h^{-2} \|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_{\Omega}^2 \max_{z \in \mathcal{N}_h} |\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z)] - \mathbf{H}[\mathbf{T}\boldsymbol{\nu}(z)]|^2$$

$$\stackrel{(4.50)}{\lesssim} \ell_{\mathrm{ex}}^2 \alpha_{\mathbf{P}}^{-1} k h^{-2} \, (\|\mathbf{x}\|_{\boldsymbol{\nu}}^{\star})^2 \max_{z \in \mathcal{N}_h} |\mathbf{H}[\mathbf{T}\boldsymbol{\mu}(z)] - \mathbf{H}[\mathbf{T}\boldsymbol{\nu}(z)]|^2.$$

This proves $\|\mathbf{x}\|_{\boldsymbol{\mu}}^{\star} \lesssim \widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})\|\mathbf{x}\|_{\boldsymbol{\nu}}^{\star}$ and hence concludes the proof. $\square$

For certain $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{M}_h$, the norm equivalence $\|\cdot\|_{\boldsymbol{\mu}}^{\star} \simeq \|\cdot\|_{\boldsymbol{\nu}}^{\star}$ holds independently of the mesh-size h.

Lemma 4.6.6. *Let $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{M}_h$ with $1 + (\mathbf{T}\boldsymbol{\mu}(z))_3 \geq \gamma > 0$ and $1 + (\mathbf{T}\boldsymbol{\nu}(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. Let*

$$\kappa(\boldsymbol{\mu}, \boldsymbol{\nu}) := \Big[1 + \gamma^{-2} \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{\infty}^2 + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_{\mathbf{P}} \gamma^2} \|\nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu}\|_{\infty}^2 \tag{4.52}$$
$$+ \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_{\mathbf{P}} \gamma^6} \big(\|\nabla \boldsymbol{\mu}\|_{\infty}^2 + \|\nabla \boldsymbol{\nu}\|_{\infty}^2 \big) \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{\infty}^2 \Big]^{1/2} \geq 1.$$

Then, $\kappa(\boldsymbol{\mu}, \boldsymbol{\nu}) = \kappa(\boldsymbol{\nu}, \boldsymbol{\mu})$ and there exists $C \geq 1$ depending only on $C_{\mathbf{mesh}}$ such that

$$C^{-1} \, \kappa(\boldsymbol{\mu}, \boldsymbol{\nu})^{-1} \, \|\mathbf{x}\|_{\boldsymbol{\nu}}^{\star} \leq \|\mathbf{x}\|_{\boldsymbol{\mu}}^{\star} \leq C \, \kappa(\boldsymbol{\mu}, \boldsymbol{\nu}) \, \|\mathbf{x}\|_{\boldsymbol{\nu}}^{\star} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

Proof. Let $\mathbf{x} \in \mathbb{R}^{2N}$. Since the symmetry $\kappa(\boldsymbol{\mu}, \boldsymbol{\nu}) = \kappa(\boldsymbol{\nu}, \boldsymbol{\mu})$ is obvious, we only have to show that $\|\mathbf{x}\|_{\boldsymbol{\mu}}^{\star} \lesssim \kappa(\boldsymbol{\mu}, \boldsymbol{\nu})\|\mathbf{x}\|_{\boldsymbol{\nu}}^{\star}$. With Lemma 4.6.1, we get that

$$\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}] = \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] + \big(\mathbb{P}_h[\boldsymbol{\mu}] - \mathbb{P}_h[\boldsymbol{\nu}] \big) \circ \mathbb{P}_h^T[\boldsymbol{\nu}] \circ \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]. \tag{4.53a}$$

With Lemma 4.6.4, Lemma 4.6.2(ii), and Lemma 4.6.3(ii), we get that

$$\|\nabla \big(\mathbb{P}_h[\boldsymbol{\mu}] - \mathbb{P}_h[\boldsymbol{\nu}] \big) \circ \mathbb{P}_h^T[\boldsymbol{\nu}] \circ \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] \, \mathbf{x}\|_{\Omega}$$

$$\stackrel{(4.42)}{\lesssim} \gamma^{-1} \|\nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu}\|_{\infty} \|\mathbb{P}_h^T[\boldsymbol{\nu}] \circ \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] \, \mathbf{x}\|_{\Omega}$$
$$+ \gamma^{-1} \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{\infty} \|\nabla \big[\mathbb{P}_h^T[\boldsymbol{\nu}] \circ \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] \, \mathbf{x} \big]\|_{\Omega}$$
$$+ \gamma^{-3} \big(\|\nabla \boldsymbol{\mu}\|_{\infty} + \|\nabla \boldsymbol{\nu}\|_{\infty} \big) \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{\infty} \|\mathbb{P}_h^T[\boldsymbol{\nu}] \circ \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] \, \mathbf{x}\|_{\Omega}$$

$$\lesssim \gamma^{-1} \|\nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu}\|_{\infty} \|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] \, \mathbf{x}\|_{\Omega} + \gamma^{-1} \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{\infty} \|\nabla \widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] \, \mathbf{x}\|_{\Omega}$$
$$+ \gamma^{-3} \big(\|\nabla \boldsymbol{\mu}\|_{\infty} + \|\nabla \boldsymbol{\nu}\|_{\infty} \big) \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_{\infty} \|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}] \, \mathbf{x}\|_{\Omega}. \tag{4.53b}$$

Recalling $\|\mathbb{P}_h[\boldsymbol{\mu}]\mathbf{x}\|_\Omega \simeq \|\mathbb{P}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega$ from Lemma 4.6.2(v), we obtain that

$$
\begin{aligned}
(\|\mathbf{x}\|_{\boldsymbol{\mu}}^\star)^2 &\overset{(4.53)}{\lesssim} \alpha_{\mathbf{P}} \|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \left[\gamma^{-2} \|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty^2 \right.\\
&\qquad + \gamma^{-6} \left(\|\nabla\boldsymbol{\mu}\|_\infty^2 + \|\nabla\boldsymbol{\nu}\|_\infty^2 \right) \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty^2 \left. \right] \|\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega^2 \\
&\qquad + \ell_{\mathrm{ex}}^2 k \left[1 + \gamma^{-2} \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty^2 \right] \|\nabla\widetilde{\mathbb{P}}_h[\boldsymbol{\nu}]\mathbf{x}\|_\Omega^2 \\
&\overset{(4.53)}{\leq} \left[1 + \gamma^{-2} \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty^2 + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_{\mathbf{P}} \gamma^2} \|\nabla\boldsymbol{\mu} - \nabla\boldsymbol{\nu}\|_\infty^2 \right. \\
&\qquad + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_{\mathbf{P}} \gamma^6} \left(\|\nabla\boldsymbol{\mu}\|_\infty^2 + \|\nabla\boldsymbol{\nu}\|_\infty^2 \right) \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty^2 \left. \right] (\|\mathbf{x}\|_{\boldsymbol{\nu}}^\star)^2.
\end{aligned}
$$

This proves $\|\mathbf{x}\|_{\boldsymbol{\mu}}^\star \lesssim \kappa(\boldsymbol{\nu}, \boldsymbol{\mu})\|\mathbf{x}\|_{\boldsymbol{\nu}}^\star$ and concludes the proof. $\qquad\square$

Lemma 4.6.7. *Let $\boldsymbol{\mu}, \boldsymbol{\nu} \in \boldsymbol{\mathcal{M}}_h$. There exists a constant $C > 1$, which depends only on $C_{\mathbf{mesh}} > 0$, such that the following two assertions (i)–(ii) hold true:*
(i) With $\widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})$ from (4.51a), it holds that, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}$,

$$
\mathbf{x} \cdot \mathbf{A_Q}[\boldsymbol{\mu}]\mathbf{x} \geq C^{-1} \frac{\alpha}{\alpha_{\mathbf{P}}} \widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})^{-2} (\|\mathbf{x}\|_{\boldsymbol{\nu}}^\star)^2, \quad \text{and} \tag{4.54a}
$$

$$
\mathbf{x} \cdot \mathbf{A_Q}[\boldsymbol{\mu}]\mathbf{y} \leq C \frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha_{\mathbf{P}}} \widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})^2 \|\mathbf{x}\|_{\boldsymbol{\nu}}^\star \|\mathbf{y}\|_{\boldsymbol{\nu}}^\star. \tag{4.54b}
$$

(ii) If, additionally, $1 + (\mathbf{T}\boldsymbol{\mu}(z))_3 \geq \gamma > 0$ and $1 + (\mathbf{T}\boldsymbol{\nu}(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, the statement of (i) holds with $\kappa(\boldsymbol{\mu}, \boldsymbol{\nu})$ from (4.52) instead of $\widetilde{\kappa}(\boldsymbol{\mu}, \boldsymbol{\nu}, h^{-2})$ from (4.51a). In particular, the estimate then is independent of the mesh-size h.

Proof. First, we prove (i). Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}$. Recall $\mathbf{A_Q}[\boldsymbol{\mu}]$ from (4.14) as well as $\mathbf{A}[\boldsymbol{\mu}], \mathbf{M}, \mathbf{L}, \mathbf{S}[\boldsymbol{\mu}]$ from Section 4.2.4. Since $\mathbf{S}[\boldsymbol{\mu}]$ is skew-symmetric, it holds that

$$
\begin{aligned}
\mathbf{x} \cdot \mathbf{A_Q}[\boldsymbol{\mu}]\mathbf{x} &\overset{(4.14)}{=} \alpha\, \mathbf{Q}[\boldsymbol{\mu}]\mathbf{x} \cdot \mathbf{MQ}[\boldsymbol{\mu}]\mathbf{x} + \ell_{\mathrm{ex}}^2 k\, \mathbf{Q}[\boldsymbol{\mu}]\mathbf{x} \cdot \mathbf{LQ}[\boldsymbol{\mu}]\mathbf{x} \\
&\overset{(4.29a)}{=} \langle \Omega, \alpha\, \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\rangle \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x} + \ell_{\mathrm{ex}}^2 k\, \langle \Omega, \nabla\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\rangle \nabla\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x} \overset{(4.50)}{\geq} \frac{\alpha}{\alpha_{\mathbf{P}}} (\|\mathbf{x}\|_{\boldsymbol{\mu}}^\star)^2.
\end{aligned}
$$

With the norm equivalence result from Lemma 4.6.5, we replace $\|\mathbf{x}\|_{\boldsymbol{\mu}}^\star$ with $\|\mathbf{x}\|_{\boldsymbol{\nu}}^\star$ and prove (4.54a). Similarly, we obtain that

$$
\begin{aligned}
\mathbf{x} \cdot \mathbf{A_Q}[\boldsymbol{\mu}]\mathbf{y} &\overset{(4.50)}{=} \langle\!\langle\!\langle \mathbf{x}, \mathbf{y}\rangle\!\rangle\!\rangle_{\boldsymbol{\mu}}^\star + \langle \Omega, (\alpha - \alpha_{\mathbf{P}})\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\rangle \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{y} + \langle \Omega, \boldsymbol{\mu} \times \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\rangle \widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{y} \\
&\leq \|\mathbf{x}\|_{\boldsymbol{\mu}}^\star \|\mathbf{y}\|_{\boldsymbol{\mu}}^\star + \left(1 + \alpha_{\mathbf{P}} - \alpha \right) \|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{x}\|_\Omega \|\widetilde{\mathbb{P}}_h[\boldsymbol{\mu}]\mathbf{y}\|_\Omega \\
&\overset{(4.50)}{\leq} \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\alpha_{\mathbf{P}} - \alpha}{\alpha_{\mathbf{P}}} \right) \|\mathbf{x}\|_{\boldsymbol{\mu}}^\star \|\mathbf{y}\|_{\boldsymbol{\mu}}^\star.
\end{aligned}
$$

Again, with the norm equivalence result from Lemma 4.6.5, we prove (4.54b). This concludes the proof of (i). The proof of (ii) follows the same lines but employs Lemma 4.6.6 instead Lemma 4.6.5. Altogether, this concludes the proof. $\qquad\square$

4.6.5 Proof of Theorem 4.4.3

The following lemma considers preconditioned GMRES with the theoretical preconditioner from (4.48).

Lemma 4.6.8. *Let $\alpha_{\mathbf{P}} \geq \alpha$ and $\boldsymbol{\mu} \in \mathcal{M}_h$. Consider the preconditioned GMRES algorithm with the preconditioner $\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{\mu}]$ from (4.48) for the solution of (4.14) with the initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^{2N}$. For $\ell \in \mathbb{N}_0$, let $\mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}$ denote the GMRES iterates with the corresponding residuals*

$$\mathbf{r}^{(\ell)} := \mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{\mu}]\mathbf{b}_{\mathbf{Q}}[m] - \mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{\mu}]\mathbf{A}_{\mathbf{Q}}[m]\,\mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}.$$

Then, there exists $0 < \kappa < 1$ such that

$$\|\mathbf{r}^{(\ell)}\|_{\boldsymbol{\mu}}^{\star} \leq (1-\kappa)^{\ell/2}\,\|\mathbf{r}^{(0)}\|_{\boldsymbol{\mu}}^{\star} \quad \text{for all } \ell \in \mathbb{N}. \tag{4.55}$$

Moreover, there hold the following assertions (i)–(ii), where $C > 1$ depends only on $C_{\mathbf{mesh}}$, and $\widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})$ as well as $\kappa(\boldsymbol{m}, \boldsymbol{\mu})$ are defined in (4.51)–(4.52):

(i) In general, κ can be chosen such that

$$\kappa \geq \left[C\left(\frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha}\right)^2 \widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})^8 \right]^{-1} > 0. \tag{4.56}$$

(ii) If $\mathbf{T}$ is chosen such that $1 + (\mathbf{T}\boldsymbol{m}(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, then κ can be chosen such that

$$\kappa \geq \left[C\left(\frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha}\right)^2 \kappa(\boldsymbol{m}, \boldsymbol{\mu})^8 \right]^{-1} > 0. \tag{4.57}$$

Proof. First, we prove (i). For a non-symmetric but positive definite system matrix, the fields-of-value analysis for the preconditioned GMRES algorithm (see, e.g., [Sta97, Theorem 3.2]) yields that

$$\|\mathbf{r}^{(\ell)}\|_{\boldsymbol{\mu}}^{\star} \leq \left(1 - \gamma^{(1)}\,\gamma^{(2)}\right)^{\ell/2}\|\mathbf{r}^{(0)}\|_{\boldsymbol{\mu}}^{\star}, \tag{4.58a}$$

where

$$\gamma^{(1)} := \inf_{\mathbf{x}\in\mathbb{R}^{2N}\setminus\{\mathbf{0}\}} \frac{\mathbf{x} \cdot \mathbf{A}_{\mathbf{Q}}[m]\mathbf{x}}{\mathbf{x} \cdot \left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{\mu}]\right)^{-1}\mathbf{x}} > 0, \tag{4.58b}$$

$$\gamma^{(2)} := \inf_{\mathbf{x}\in\mathbb{R}^{2N}\setminus\{\mathbf{0}\}} \frac{\mathbf{x} \cdot \left(\mathbf{A}_{\mathbf{Q}}[m]\right)^{-1}\mathbf{x}}{\mathbf{x} \cdot \mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{\mu}]\mathbf{x}} > 0. \tag{4.58c}$$

To estimate $\gamma^{(1)}, \gamma^{(2)}$ from below, recall $\widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})$ from (4.51a) and exploit Lemma 4.6.7(i). This yields that

$$\gamma^{(1)} \overset{(4.51)}{\gtrsim} \frac{\alpha}{\alpha_{\mathbf{P}}}\widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})^{-2}.$$

With Lemma 4.6.7(i), the matrices $\mathbf{B} := \mathbf{A_Q}[m]$ and $\mathbf{B}_0 := \left(\mathbf{P_Q^\star}[\boldsymbol{\mu}]\right)^{-1}$ satisfy the setting of Lemma 4.6.11 with

$$c_1 \simeq \frac{\alpha}{\alpha_{\mathbf{P}}} \widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})^{-2} \quad \text{and} \quad c_2 \simeq \frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha_{\mathbf{P}}} \widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})^2.$$

Hence, Lemma 4.6.11 yields that

$$\gamma^{(2)} \overset{(4.58c)}{\gtrsim} \frac{c_1}{c_2^2} \simeq \frac{\alpha}{\alpha_{\mathbf{P}}} \left(\frac{2\alpha_{\mathbf{P}} + 1 - \alpha}{\alpha_{\mathbf{P}}} \right)^{-2} \widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})^{-6}. \tag{4.59}$$

With (4.58), we conclude the proof of (i). The proof of (ii) then follows the same lines but exploits Lemma 4.6.7(ii) instead of Lemma 4.6.7(i). In the latter arguments, this replaces $\widetilde{\kappa}(\boldsymbol{m}, \boldsymbol{\mu}, h^{-2})$ by $\kappa(\boldsymbol{m}, \boldsymbol{\mu})$. Altogether, this concludes the proof. □

Proof of Theorem 4.4.3. Recall that $\mathbf{T} = \mathbf{T}^{-1} = \mathbf{T}^T$. For constant $\boldsymbol{\mu} := \mathbf{T}\mathbf{e}_3 \in \mathcal{M}_h$, we get that

$$\mathbf{TH}[\mathbf{T}\boldsymbol{\mu}(z_i)] = \mathbf{TH}[\mathbf{e}_3] = \mathbf{T}[\mathbf{e}_1, \mathbf{e}_2] \in \mathbb{R}^{3 \times 2} \tag{4.60a}$$

and thus

$$(\mathbf{TH}[\mathbf{T}\boldsymbol{\mu}(z_i)])^T (\mathbf{TH}[\mathbf{T}\boldsymbol{\mu}(z_j)]) = [\mathbf{e}_1, \mathbf{e}_2]^T \mathbf{TT}[\mathbf{e}_1, \mathbf{e}_2] = \mathbf{I}_{2 \times 2}. \tag{4.60b}$$

for all $i, j = 1, \ldots, N$. Together with the block forms from (4.8), this yields that

$$\mathbf{P_Q^\star}[\boldsymbol{\mu}] \overset{(4.48)}{=} \left(\mathbf{Q}[\boldsymbol{\mu}]^T \left(\alpha_{\mathbf{P}} \mathbf{M} + \ell_{\mathrm{ex}}^2 k \mathbf{L} \right) \mathbf{Q}[\boldsymbol{\mu}] \right)^{-1} \overset{(4.60)}{=} \left(\alpha_{\mathbf{P}} \mathbf{M}_{\mathrm{2D}} + \ell_{\mathrm{ex}}^2 k \mathbf{L}_{\mathrm{2D}} \right)^{-1} \overset{(4.26)}{=} \mathbf{P}_{\mathrm{2D}}.$$

Since $\nabla \boldsymbol{\mu} = \mathbf{0}$ a.e. in Ω and $\|\boldsymbol{m} - \boldsymbol{\mu}\|_\infty \le 2$, Lemma 4.6.8(ii) proves the result. □

4.6.6 Proof of Theorem 4.4.1

In analogy to (4.58), the fields-of-value analysis for the preconditioned GMRES algorithm (see, e.g., [Sta97, Theorem 3.2]) yields that

$$\|\mathbf{r}^{(\ell)}\|_{\boldsymbol{m}} \le \left(1 - \widetilde{\gamma}^{(1)} \widetilde{\gamma}^{(2)} \right)^{\ell/2} \|\mathbf{r}^{(0)}\|_{\boldsymbol{m}}, \tag{4.61a}$$

where

$$\widetilde{\gamma}^{(1)} := \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \mathbf{A_Q}[m]\mathbf{x}}{\mathbf{x} \cdot \left(\mathbf{P_Q}[m] \right)^{-1} \mathbf{x}} > 0, \tag{4.61b}$$

$$\widetilde{\gamma}^{(2)} := \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \left(\mathbf{A_Q}[m] \right)^{-1} \mathbf{x}}{\mathbf{x} \cdot \mathbf{P_Q}[m]\mathbf{x}} > 0. \tag{4.61c}$$

Recall from (4.48), the definition of the theoretical preconditioner $\mathbf{P_Q^\star}[m]$ and from (4.58) the corresponding definition of $\gamma^{(1)}$ and $\gamma^{(2)}$. We obtain that

$$\widetilde{\gamma}^{(1)} \overset{(4.61b)}{\ge} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \mathbf{A_Q}[m]\mathbf{x}}{\mathbf{x} \cdot \left(\mathbf{P_Q^\star}[m] \right)^{-1} \mathbf{x}} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \left(\mathbf{P_Q^\star}[m] \right)^{-1} \mathbf{x}}{\mathbf{x} \cdot \left(\mathbf{P_Q}[m] \right)^{-1} \mathbf{x}}$$

$$\overset{(4.58b)}{=} \gamma^{(1)} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \left(\mathbf{P_Q^\star}[m] \right)^{-1} \mathbf{x}}{\mathbf{x} \cdot \left(\mathbf{P_Q}[m] \right)^{-1} \mathbf{x}} =: \gamma^{(1)} \delta^{(1)}, \tag{4.62a}$$

as well as

$$\widetilde{\gamma}^{(2)} \overset{(4.61b)}{\geq} \inf_{\mathbf{x}\in\mathbb{R}^{2N}\setminus\{\mathbf{0}\}} \frac{\mathbf{x}\cdot\left(\mathbf{A}_{\mathbf{Q}}[\boldsymbol{m}]\right)^{-1}\mathbf{x}}{\mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\,\mathbf{x}} \inf_{\mathbf{x}\in\mathbb{R}^{2N}\setminus\{\mathbf{0}\}} \frac{\mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\mathbf{x}}{\mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}[\boldsymbol{m}]\mathbf{x}}$$

$$\overset{(4.58b)}{=} \gamma^{(2)} \inf_{\mathbf{x}\in\mathbb{R}^{2N}\setminus\{\mathbf{0}\}} \frac{\mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\,\mathbf{x}}{\mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}[\boldsymbol{m}]\mathbf{x}} =: \gamma^{(2)}\,\delta^{(2)}. \tag{4.62b}$$

Here, we implicitly have $\boldsymbol{m} = \boldsymbol{\mu}$ in the definition (4.58) of $\gamma^{(1)}$ and $\gamma^{(2)}$. In particular, Lemma 4.6.7 holds with $\widetilde{\kappa}(\boldsymbol{m},\boldsymbol{m},h^{-2}) = \kappa(\boldsymbol{m},\boldsymbol{m}) = 1$. Following the lines of the proof of Lemma 4.6.8, this yields that

$$\gamma^{(1)} \gtrsim \frac{\alpha}{\alpha_{\mathbf{P}}}, \quad \text{and} \quad \gamma^{(2)} \overset{(4.59)}{\gtrsim} \frac{\alpha}{\alpha_{\mathbf{P}}}\left(\frac{2\alpha_{\mathbf{P}}+1-\alpha}{\alpha_{\mathbf{P}}}\right)^{-2}. \tag{4.63}$$

Hence, in the following four steps, it remains to estimate $\delta^{(1)}$ and $\delta^{(2)}$ from below.

Step 1. We will use the fictitious space lemma (see [Nep91, GO95]) to derive

$$c_1\,\mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\mathbf{x} \leq \mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}[\boldsymbol{m}]\mathbf{x} \leq c_2\,\mathbf{x}\cdot\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\mathbf{x} \quad \text{for all } \mathbf{x}\in\mathbb{R}^{2N}. \tag{4.64}$$

Here, the constants $c_1, c_2 > 0$ stem from the following two assumptions **(FS1)**–**(FS2)** of the fictitious space lemma:

(FS1) For all $\mathbf{x}\in\mathbb{R}^{2N}$, there exists $\mathbf{y}\in\mathbb{R}^{3N}$ with $\mathbf{Q}[\boldsymbol{m}]^T\mathbf{y} = \mathbf{x}$ and

$$c_1\,\mathbf{y}\cdot(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\mathrm{ex}}^2 k\,\mathbf{L})\mathbf{y} \leq \mathbf{x}\cdot\left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\mathbf{x}. \tag{4.65a}$$

(FS2) For all $\mathbf{y}\in\mathbb{R}^{3N}$, it holds that

$$\mathbf{Q}[\boldsymbol{m}]^T\mathbf{y}\cdot\left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\mathbf{Q}[\boldsymbol{m}]^T\mathbf{y} \leq c_2\,\mathbf{y}\cdot(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\mathrm{ex}}^2 k\,\mathbf{L})\mathbf{y}. \tag{4.65b}$$

With the assumptions **(FS1)**–**(FS2)**, the fictitious space lemma then implies that

$$c_1\,\widetilde{\mathbf{x}}\cdot\left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\widetilde{\mathbf{x}} \;\leq\; \left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\widetilde{\mathbf{x}}\cdot\mathbf{Q}[\boldsymbol{m}]^T(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\mathrm{ex}}^2 k\,\mathbf{L})^{-1}\mathbf{Q}[\boldsymbol{m}]\left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\widetilde{\mathbf{x}}$$

$$\overset{(4.16)}{=} \left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\widetilde{\mathbf{x}}\cdot\mathbf{P}_{\mathbf{Q}}[\boldsymbol{m}]\left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\widetilde{\mathbf{x}}$$

$$\leq\; c_2\,\widetilde{\mathbf{x}}\cdot\left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\widetilde{\mathbf{x}} \quad \text{for all } \widetilde{\mathbf{x}}\in\mathbb{R}^{2N}.$$

Then, with $\widetilde{\mathbf{x}} := \mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\mathbf{x}\in\mathbb{R}^{2N}$ in the latter estimate, this verifies (4.64).

Step 2. We verify assumption **(FS1)** of the fictitious space lemma. To that end, let $\mathbf{x}\in\mathbb{R}^{2N}$ and set $\mathbf{y} := \mathbf{Q}[\boldsymbol{m}]\mathbf{x}\in\mathbb{R}^{3N}$. Then, $\mathbf{Q}[\boldsymbol{m}]^T\mathbf{y} = \mathbf{Q}[\boldsymbol{m}]^T\mathbf{Q}[\boldsymbol{m}]\mathbf{x} = \mathbf{x}$ and

$$\mathbf{y}\cdot(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\mathrm{ex}}^2 k\,\mathbf{L})\mathbf{y} \;=\; \mathbf{Q}[\boldsymbol{m}]\mathbf{x}\cdot(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\mathrm{ex}}^2 k\,\mathbf{L})\mathbf{Q}[\boldsymbol{m}]\mathbf{x}$$

$$=\; \mathbf{x}\cdot\mathbf{Q}[\boldsymbol{m}]^T(\alpha_{\mathbf{P}}\mathbf{M} + \ell_{\mathrm{ex}}^2 k\,\mathbf{L})\mathbf{Q}[\boldsymbol{m}]\mathbf{x} \overset{(4.48)}{=} \mathbf{x}\cdot\left(\mathbf{P}_{\mathbf{Q}}^{\star}[\boldsymbol{m}]\right)^{-1}\mathbf{x},$$

i.e., assumption **(FS1)** holds with $c_1 = 1$.

Step 3. We verify assumption **(FS2)** of the fictitious space lemma. To that end, let $\mathbf{y} \in \mathbb{R}^{3N}$. Define $\boldsymbol{v} := \sum_{i=1}^{3N} \mathbf{y}_i \boldsymbol{\phi}_i \in \boldsymbol{\mathcal{S}}_h$. With Lemma 4.6.1(ii), we obtain that

$$
\mathbf{Q}[\boldsymbol{m}]^T \mathbf{y} \cdot \left(\mathbf{P}_\mathbf{Q}^\star[\boldsymbol{m}] \right)^{-1} \mathbf{Q}[\boldsymbol{m}]^T \mathbf{y} \overset{(4.48)}{=} \mathbf{y} \cdot \mathbf{Q}[\boldsymbol{m}] \mathbf{Q}[\boldsymbol{m}]^T (\alpha_\mathbf{P} \mathbf{M} + \ell_{\mathrm{ex}}^2 k \, \mathbf{L}) \mathbf{Q}[\boldsymbol{m}] \mathbf{Q}[\boldsymbol{m}]^T \mathbf{y}
$$
$$
\overset{(4.31b)}{=} \alpha_\mathbf{P} \| \widetilde{\boldsymbol{\Pi}}_h[\boldsymbol{m}] \mathbf{y} \|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \, \| \nabla \widetilde{\boldsymbol{\Pi}}_h[\boldsymbol{m}] \mathbf{y} \|_\Omega^2
$$
$$
= \alpha_\mathbf{P} \| \boldsymbol{\Pi}_h[\boldsymbol{m}] \boldsymbol{v} \|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \, \| \nabla \boldsymbol{\Pi}_h[\boldsymbol{m}] \boldsymbol{v} \|_\Omega^2. \qquad (4.66)
$$

For the verification of **(FS2)** in (i), Lemma 4.6.2(iii) and an inverse estimate yield that

$$
\alpha_\mathbf{P} \| \boldsymbol{\Pi}_h[\boldsymbol{m}] \boldsymbol{v} \|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \, \| \nabla \boldsymbol{\Pi}_h[\boldsymbol{m}] \boldsymbol{v} \|_\Omega^2 \lesssim \left(\alpha_\mathbf{P} + \frac{\ell_{\mathrm{ex}}^2 k}{h^2} \right) \| \boldsymbol{v} \|_\Omega^2
$$
$$
\lesssim \left(1 + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_\mathbf{P} h^2} \right) \left[\alpha_\mathbf{P} \| \boldsymbol{v} \|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \, \| \nabla \boldsymbol{v} \|_\Omega^2 \right] = \left(1 + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_\mathbf{P} h^2} \right) \mathbf{y} \cdot \left(\alpha_\mathbf{P} \mathbf{M} + \ell_{\mathrm{ex}}^2 k \, \mathbf{L} \right) \mathbf{y}.
$$

$$(4.67a)$$

For the verification of **(FS2)** in (ii), we use the stronger assumption $1 + (\mathbf{T}\boldsymbol{m})_3 \geq \gamma > 0$. Then, the definition of $\boldsymbol{v}$, Lemma 4.6.2(iii) and Lemma 4.6.3(iii) yield that

$$
\alpha_\mathbf{P} \| \boldsymbol{\Pi}_h[\boldsymbol{m}] \boldsymbol{v} \|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \, \| \nabla \boldsymbol{\Pi}_h[\boldsymbol{m}] \boldsymbol{v} \|_\Omega^2 \lesssim \left(\alpha_\mathbf{P} + \frac{\ell_{\mathrm{ex}}^2 k}{\gamma^4} \| \nabla \boldsymbol{m} \|_\infty^2 \right) \| \boldsymbol{v} \|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \| \nabla \boldsymbol{v} \|_\Omega^2
$$
$$
\leq \left(1 + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_\mathbf{P} \gamma^4} \| \nabla \boldsymbol{m} \|_\infty^2 \right) \left[\alpha_\mathbf{P} \| \boldsymbol{v} \|_\Omega^2 + \ell_{\mathrm{ex}}^2 k \, \| \nabla \boldsymbol{v} \|_\Omega^2 \right]
$$
$$
= \left(1 + \frac{\ell_{\mathrm{ex}}^2 k}{\alpha_\mathbf{P} \gamma^4} \| \nabla \boldsymbol{m} \|_\infty^2 \right) \mathbf{y} \cdot \left(\alpha_\mathbf{P} \mathbf{M} + \ell_{\mathrm{ex}}^2 k \, \mathbf{L} \right) \mathbf{y}. \qquad (4.67b)
$$

We combine (4.66)–(4.67) and obtain that **(FS2)** holds with

$$
c_2 \lesssim \begin{cases} 1 + \dfrac{\ell_{\mathrm{ex}}^2 k}{\alpha_\mathbf{P} \gamma^4} \| \nabla \boldsymbol{m} \|_\infty^2 & \text{if } 1 + (\mathbf{T}\boldsymbol{m})_3 \geq \gamma > 0, \\ 1 + \dfrac{\ell_{\mathrm{ex}}^2 k}{\alpha_\mathbf{P} h^2} & \text{else.} \end{cases} \qquad (4.68)
$$

Step 4. With **Step 1**–**Step 3**, the matrices $\mathbf{B} := \mathbf{P}_\mathbf{Q}[\boldsymbol{m}]$ and $\mathbf{B}_0 := \mathbf{P}_\mathbf{Q}^\star[\boldsymbol{m}]$ satisfy the assumptions of Lemma 4.6.11 with $c_1 = 1$ and c_2 from (4.68). Hence, we get that

$$
\frac{1}{c_2^2} \mathbf{x} \cdot \left(\mathbf{P}_\mathbf{Q}^\star[\boldsymbol{m}] \right)^{-1} \mathbf{x} \leq \mathbf{x} \cdot \left(\mathbf{P}_\mathbf{Q}[\boldsymbol{m}] \right)^{-1} \mathbf{x} \leq \mathbf{x} \cdot \left(\mathbf{P}_\mathbf{Q}^\star[\boldsymbol{m}] \right)^{-1} \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}. \qquad (4.69)
$$

From this and (4.64), we obtain that

$$
\delta^{(1)} \overset{(4.62a)}{=} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \left(\mathbf{P}_\mathbf{Q}^\star[\boldsymbol{m}] \right)^{-1} \mathbf{x}}{\mathbf{x} \cdot \left(\mathbf{P}_\mathbf{Q}[\boldsymbol{m}] \right)^{-1} \mathbf{x}} \overset{(4.69)}{\geq} 1 \quad \text{and} \quad \delta^{(2)} \overset{(4.62b)}{=} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \mathbf{P}_\mathbf{Q}^\star[\boldsymbol{m}] \mathbf{x}}{\mathbf{x} \cdot \mathbf{P}_\mathbf{Q}[\boldsymbol{m}] \mathbf{x}} \overset{(4.64)}{\geq} \frac{1}{c_2}.
$$

Together with the estimates for $\gamma^{(1)}$ and $\gamma^{(2)}$ from (4.63), this concludes the proof. $\qquad \square$

4.6.7 Auxiliary results

This section collects some technical lemmas used in the proofs of the main results.

Lemma 4.6.9. *For $(\mu_1, \mu_2, \mu_3)^T \equiv \boldsymbol{\mu} \in \boldsymbol{\mathcal{M}}_h$, there hold the following assertions (i)–(iii):*
(i) For all $i, j \in \{1, 2\}$, it holds that

$$\left\| \frac{\mu_i \mu_j}{1 + \mu_3} \right\|_\infty \leq 2.$$

(ii) Let $1 + \mu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. For all $i, j \in \{1, 2\}$, it holds that

$$\frac{\mu_i \mu_j}{1 + \mu_3} \in W^{1,\infty}(\Omega) \quad \text{with} \quad \left\| \partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] \right\|_\infty \leq 3\,\gamma^{-1} \|\partial_k \boldsymbol{\mu}\|_\infty$$

for all $k \in \{1, 2, 3\}$.
(iii) Let $1 + \mu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. For all $i, j \in \{1, 2\}$, it holds that

$$\frac{\mu_i \mu_j}{1 + \mu_3} \in W^{2,\infty}(K) \quad \text{with} \quad \left\| \partial_\ell \partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] \right\|_\infty \leq 12\,\gamma^{-2} \|\partial_k \boldsymbol{\mu}\|_\infty \|\partial_\ell \boldsymbol{\mu}\|_\infty,$$

for all elements $K \in \mathcal{T}_h$ and for all $\ell, k \in \{1, 2, 3\}$.

Proof. To prove (i), note that piecewise affine functions attain their maximal length at the nodes. Since $\boldsymbol{\mu} \in \boldsymbol{\mathcal{M}}_h$, together with Young's inequality this yield that

$$\left| \frac{\mu_i \mu_j}{1 + \mu_3} \right| \leq \frac{\mu_1^2 + \mu_2^2}{1 + \mu_3} \leq \frac{1 - \mu_3^2}{1 + \mu_3} = \frac{(1 - \mu_3)(1 + \mu_3)}{1 + \mu_3} = (1 - \mu_3) \leq 2.$$

This proves (i). For the proof of (ii) and (iii), the product rule yields that

$$\partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] = \frac{(\partial_k \mu_i) \mu_j}{1 + \mu_3} + \frac{\mu_i (\partial_k \mu_j)}{1 + \mu_3} - \frac{\mu_i \mu_j (\partial_k \mu_3)}{(1 + \mu_3)^2}. \tag{4.70}$$

Moreover, we exploit that the second derivative of affine functions is zero. Elementwise, the product rule then yields that

$$\partial_\ell \partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] \overset{(4.70)}{=} \frac{(\partial_k \mu_i)(\partial_\ell \mu_j)}{1 + \mu_3} - \frac{(\partial_k \mu_i) \mu_j (\partial_\ell \mu_3)}{(1 + \mu_3)^2} + \frac{(\partial_\ell \mu_i)(\partial_k \mu_j)}{1 + \mu_3} - \frac{\mu_i (\partial_k \mu_j)(\partial_\ell \mu_3)}{(1 + \mu_3)^2}$$
$$- \frac{(\partial_\ell \mu_i) \mu_j (\partial_k \mu_3)}{(1 + \mu_3)^2} - \frac{\mu_i (\partial_\ell \mu_j)(\partial_k \mu_3)}{(1 + \mu_3)^2} + \frac{2 \mu_i \mu_j (\partial_k \mu_3)(\partial_\ell \mu_3)}{(1 + \mu_3)^3}.$$
$$\tag{4.71}$$

Since $|\boldsymbol{\mu}| \leq 1$ and since $\boldsymbol{\mu}$ is piecewise affine, the assumption $1 + \mu_3(z) \geq \gamma > 0$ for all nodes implies that $1 + \mu_3 \geq \gamma > 0$ in Ω. Together with $|\boldsymbol{\mu}| \leq 1$ and (4.70), this proves (ii). For the proof of (iii), additionally note that $\gamma \leq 1 + \mu_3 \leq 1 + |\mu_3| \leq 2$ yields that $1 \leq 2/\gamma$. Together with $|\boldsymbol{\mu}| \leq 1$ and (4.71), this proves (iii). $\qquad \square$

Lemma 4.6.10. *For $(\mu_1, \mu_2, \mu_3)^T \equiv \boldsymbol{\mu}, (\nu_1, \nu_2, \nu_3)^T \equiv \boldsymbol{\nu} \in \mathcal{M}_h$, let $1 + \mu_3(z) \geq \gamma > 0$ and $1 + \nu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. Then, there exists a constant $C > 0$ such that the following assertions (i)–(iii) hold true:*

(i) For all $i, j \in \{1, 2\}$, it holds that

$$\left\| \frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} \right\|_\infty \leq C \gamma^{-1} \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty.$$

(ii) For all $i, j \in \{1, 2\}$ and all $k \in \{1, 2, 3\}$, it holds that

$$\left\| \partial_k \left(\frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} \right) \right\|_\infty \leq C \gamma^{-2} \left(\|\nabla \boldsymbol{\mu}\|_\infty + \|\nabla \boldsymbol{\nu}\|_\infty \right) \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty \\ + C \gamma^{-1} \|\nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu}\|_\infty.$$

(iii) For all $i, j \in \{1, 2\}$, all elements $K \in \mathcal{T}_h$ and all $\ell, k \in \{1, 2, 3\}$, it holds that

$$\left\| \partial_\ell \partial_k \left(\frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} \right) \right\|_\infty \leq C \gamma^{-3} \left(\|\nabla \boldsymbol{\mu}\|_\infty + \|\nabla \boldsymbol{\nu}\|_\infty \right)^2 \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty \\ + C \gamma^{-2} \left(\|\nabla \boldsymbol{\mu}\|_\infty + \|\nabla \boldsymbol{\nu}\|_\infty \right) \|\nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu}\|_\infty.$$

Proof. Throughout the proof, we write

$$P_{ij} := \frac{\mu_i \mu_j}{1 + \mu_3} \quad \text{and} \quad d_k := \mu_k - \nu_k \quad \text{for all } i, j \in \{1, 2\} \text{ and all } k \in \{1, 2, 3\}. \tag{4.72}$$

Recall that the properties of P_{ij} are discussed in Lemma 4.6.9. Since $\boldsymbol{\mu}, \boldsymbol{\nu}$ are piecewise affine, we get that $|\boldsymbol{\mu}|, |\boldsymbol{\nu}| \leq 1$ on Ω. Moreover, since $1 + \mu_3(z) \geq \gamma > 0$ and $1 + \nu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, it follows that $1 + \mu_3 \geq \gamma > 0$ and $1 + \nu_3 \geq \gamma > 0$ on Ω. For the proof of (i), elementary computations show that

$$\frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} = -P_{ij} \frac{d_3}{1 + \nu_3} + \frac{\mu_j d_i}{1 + \nu_3} + \frac{\nu_i d_j}{1 + \nu_3}. \tag{4.73}$$

Together with Lemma 4.6.9 (i), this proves (i). For the proof of (ii), let $k \in \{1, 2, 3\}$. We differentiate the terms in (4.73) separately and obtain that

$$\partial_k \left(P_{ij} \frac{d_3}{1 + \nu_3} \right) = (\partial_k P_{ij}) \frac{d_3}{1 + \nu_3} + P_{ij} \frac{(\partial_k d_3)}{1 + \nu_3} - P_{ij} \frac{d_3 (\partial_k \nu_3)}{(1 + \nu_3)^2} =: T_1 + T_2 + T_3, \tag{4.74a}$$

$$\partial_k \left(\frac{\mu_j d_i}{1 + \nu_3} \right) = \frac{(\partial_k \mu_j) d_i}{1 + \nu_3} + \frac{\mu_j (\partial_k d_i)}{1 + \nu_3} - \frac{\mu_j d_i (\partial_k \nu_3)}{(1 + \nu_3)^2} =: T_4 + T_5 + T_6, \tag{4.74b}$$

$$\partial_k \left(\frac{\nu_i d_j}{1 + \nu_3} \right) = \frac{(\partial_k \nu_i) d_j}{1 + \nu_3} + \frac{\nu_i (\partial_k d_j)}{1 + \nu_3} - \frac{\nu_i d_j (\partial_k \nu_3)}{(1 + \nu_3)^2} =: T_7 + T_8 + T_9. \tag{4.74c}$$

With $1 + \mu_3 \geq \gamma > 0$ and $1 + \nu_3 \geq \gamma > 0$, Lemma 4.6.9 (i)–(ii) yields that

$$\sum_{i \in \{1,3,4,6,7,9\}} |T_i| \lesssim \gamma^{-2} \left(\|\nabla \boldsymbol{\mu}\|_\infty + \|\nabla \boldsymbol{\nu}\|_\infty \right) \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty, \quad \text{and}$$

$$\sum_{i \in \{2,5,8\}} |T_i| \lesssim \gamma^{-1} \|\nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu}\|_\infty.$$

Together with (4.73), this proves (ii). For the proof of (iii), let $\ell, k \in \{1, 2, 3\}$. We differentiate the terms in (4.74) separately and exploit that the second derivative of piecewise affine functions is zero. We start from (4.74a). Elementwise, the product rule yields that

$$\partial_\ell \left((\partial_k P_{ij}) \frac{d_3}{1 + \nu_3} \right) = (\partial_\ell \partial_k P_{ij}) \frac{d_3}{1 + \nu_3} + (\partial_k P_{ij}) \frac{\partial_\ell d_3}{1 + \nu_3} - (\partial_k P_{ij}) \frac{d_3 (\partial_\ell \nu_3)}{(1 + \nu_3)^2},$$
$$=: \widetilde{T}_1 + \widetilde{T}_2 + \widetilde{T}_3. \tag{4.75a}$$

$$\partial_\ell \left(P_{ij} \frac{(\partial_k d_3)}{1 + \nu_3} \right) = (\partial_\ell P_{ij}) \frac{(\partial_k d_3)}{1 + \nu_3} - P_{ij} \frac{(\partial_k d_3)(\partial_\ell \nu_3)}{(1 + \nu_3)^2} =: \widetilde{T}_4 + \widetilde{T}_5, \tag{4.75b}$$

$$\partial_\ell \left(P_{ij} \frac{d_3 (\partial_k \nu_3)}{(1 + \nu_3)^2} \right) = (\partial_\ell P_{ij}) \frac{d_3 (\partial_k \nu_3)}{(1 + \nu_3)^2} + P_{ij} \frac{(\partial_\ell d_3)(\partial_k \nu_3)}{(1 + \nu_3)^2} - 2 P_{ij} \frac{d_3 (\partial_k \nu_3)(\partial_\ell \nu_3)}{(1 + \nu_3)^3}$$
$$=: \widetilde{T}_6 + \widetilde{T}_7 + \widetilde{T}_8. \tag{4.75c}$$

Next, we get for the terms from (4.74b) elementwise that

$$\partial_\ell \left(\frac{(\partial_k \mu_j) d_i}{1 + \nu_3} \right) = \frac{(\partial_k \mu_j)(\partial_\ell d_i)}{1 + \nu_3} - \frac{(\partial_k \mu_j) d_i (\partial_\ell \nu_3)}{(1 + \nu_3)^2} =: \widetilde{T}_9 + \widetilde{T}_{10}, \tag{4.76a}$$

$$\partial_\ell \left(\frac{\mu_j (\partial_k d_i)}{1 + \nu_3} \right) = \frac{(\partial_\ell \mu_j)(\partial_k d_i)}{1 + \nu_3} - \frac{\mu_j (\partial_k d_i)(\partial_\ell \nu_3)}{(1 + \nu_3)^2} =: \widetilde{T}_{11} + \widetilde{T}_{12}, \tag{4.76b}$$

$$\partial_\ell \left(\frac{\mu_j d_i (\partial_k \nu_3)}{(1 + \nu_3)^2} \right) = \frac{(\partial_\ell \mu_j) d_i (\partial_k \nu_3)}{(1 + \nu_3)^2} + \frac{\mu_j (\partial_\ell d_i)(\partial_k \nu_3)}{(1 + \nu_3)^2} - 2 \frac{\mu_j d_i (\partial_k \nu_3)(\partial_\ell \nu_3)}{(1 + \nu_3)^3}$$
$$=: \widetilde{T}_{13} + \widetilde{T}_{14} + \widetilde{T}_{15}. \tag{4.76c}$$

Lemma 4.6.9, $1 \le 2/\gamma$ and $1 + \mu_3 \ge \gamma > 0$ as well as $1 + \nu_3 \ge \gamma > 0$ yield that

$$\sum_{i \in \{1,3,6,8,10,12,13,15\}} |\widetilde{T}_i| \lesssim \gamma^{-3} \left(\|\nabla \boldsymbol{\mu}\|_\infty + \|\nabla \boldsymbol{\nu}\|_\infty \right)^2 \|\boldsymbol{\mu} - \boldsymbol{\nu}\|_\infty, \quad \text{and}$$

$$\sum_{i \in \{2,4,5,7,9,11,14\}} |\widetilde{T}_i| \lesssim \gamma^{-2} \left(\|\nabla \boldsymbol{\mu}\|_\infty + \|\nabla \boldsymbol{\nu}\|_\infty \right) \|\nabla \boldsymbol{\mu} - \nabla \boldsymbol{\nu}\|_\infty.$$

Note that the terms in (4.74c) are obtained if we replace μ_i with ν_i and d_i with d_j in (4.74b). Hence, we can apply the same arguments as in (4.76). This proves (iii). □

Lemma 4.6.11 ([ABV14, Lemma 3.1]). *Let* $\mathbf{B} \in \mathbb{R}^{2N \times 2N}$ *be a positive definite matrix and* $\mathbf{B}_0 \in \mathbb{R}^{2N \times 2N}$ *be a symmetric positive definite matrix, which satisfy for* $c_1, c_2 > 0$ *that*

$$\mathbf{x} \cdot \mathbf{B}\mathbf{x} \ge c_1 \, \mathbf{x} \cdot \mathbf{B}_0 \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N} \quad \text{and}$$
$$\mathbf{x} \cdot \mathbf{B}\mathbf{y} \le c_2 \left(\mathbf{x} \cdot \mathbf{B}_0 \mathbf{x} \right)^{1/2} \left(\mathbf{y} \cdot \mathbf{B}_0 \mathbf{y} \right)^{1/2} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}.$$

Then, it holds that

$$\mathbf{x} \cdot \mathbf{B}^{-1} \mathbf{x} \ge \frac{c_1}{c_2^2} \, \mathbf{x} \cdot \mathbf{B}_0^{-1} \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N} \quad \text{and}$$
$$\mathbf{x} \cdot \mathbf{B}^{-1} \mathbf{y} \le c_1^{-1} \left(\mathbf{x} \cdot \mathbf{B}_0^{-1} \mathbf{x} \right)^{1/2} \left(\mathbf{y} \cdot \mathbf{B}_0^{-1} \mathbf{y} \right)^{1/2} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}. \qquad \square$$

4.7 Conclusion

We consider the solution of the constrained linear system which arises from the discretization of LLG with tangent plane schemes. For each time-step, the method requires solving for the vector-valued discrete time derivative, defined by its N nodal values in $\mathbb{R}^3$, where $N \in \mathbb{N}$ denotes the number of nodes of the mesh used for the spatial discretization. The time derivative is defined by the variational formulation posed in the time-dependent discrete tangent space of dimension $2N$. Naive saddle point approaches consider the variational formulation in the vector-valued standard finite element space $(\mathcal{S}_h)^3$. The resulting linear system is supplemented with N linear constraints which impose orthogonality between the magnetization and its time derivative, and hence leading to an indefinite system in $\mathbb{R}^{4N}$. In this work, we design a solution strategy, based on local Householder reflections, which results in a linear positive-definite system in $\mathbb{R}^{2N}$, thus recovering the effective dimension of the problem. For the proposed approach, we discuss the iterative solution of the system with the GMRES algorithm. Since applying the GMRES algorithm without preconditioner potentially leads to high iteration numbers, we derive three types of preconditioning strategies, which we refer to as practical preconditioner $\mathbf{P_Q}[\cdot]$, stationary preconditioner $\mathbf{P}_{2D}$, and Jacobi-type preconditioner $\mathbf{P}^{jac}$. For the first two strategies, we establish results which provide explicit bounds for the number of iterations required to reach a prescribed tolerance (Theorem 4.4.1 and Theorem 4.4.3, respectively). Then, we numerically compare the three proposed approaches.

In all performed experiments, using the Jacobi-type preconditioner $\mathbf{P}^{jac}$ reduces the iteration numbers; see Figure 4.1–4.2, Figure 4.6 and Figure 4.8. Since applying $\mathbf{P}^{jac}$ requires only the (trivial) inversion of a diagonal matrix, the preconditioned GMRES algorithm with $\mathbf{P}^{jac}$ should always be preferred over no preconditioning.

However, as shown in Figure 4.1 and Figure 4.8, $\mathbf{P}^{jac}$ is not robust with respect to the mesh-size. The experiments show that on fine meshes one should use either the stationary preconditioner $\mathbf{P}_{2D}$ or the practical preconditioner $\mathbf{P_Q}[\cdot]$. Based on the experiments of Section 4.5.2.1 and Section 4.5.3, we recommend choosing $\alpha_\mathbf{P} = 1$ for the preconditioners rather than the canonical choice $\alpha_\mathbf{P} = \alpha$. In Figure 4.1 and Figure 4.8, unlike $\mathbf{P}^{jac}$, these two preconditioners are shown to be robust with respect to the mesh-size. If $\mathbf{T}$ is chosen accordingly, the robustness with respect to the mesh-size can also be proved rigorously; see Theorem 4.4.1(ii) and Theorem 4.4.3 as well as Remark 4.4.2(ii).

For almost all experiments the iteration numbers for $\mathbf{P_Q}[\cdot]$ are slightly, but insignificantly lower than those for $\mathbf{P}_{2D}$. We conclude by pointing out two aspects when choosing between $\mathbf{P}_{2D}$ and $\mathbf{P_Q}[\cdot]$: While we obtain satisfying results (Figure 4.5a and Figure 4.6) for $\mathbf{P}_{2D}$ when choosing $\mathbf{T}$ adaptively as described in Section 4.4.2, the choice of $\mathbf{T}$ seems to be irrelevant for the performance of $\mathbf{P_Q}[\cdot]$; see Figure 4.5b. However, one drawback of $\mathbf{P_Q}[\cdot]$ compared to $\mathbf{P}_{2D}$ is that its application is roughly 50% more costly: This is due to the fact that in each GMRES iteration, employing $\mathbf{P_Q}[\cdot]$ essentially requires applying the inverse of $(\alpha_\mathbf{P}\mathbf{M} + \ell_{ex}^2 k\mathbf{L}) \in \mathbb{R}^{3N \times 3N}$ to a vector in $\mathbb{R}^{3N}$, while employing $\mathbf{P}_{2D}$ requires applying the inverse of $(\alpha_\mathbf{P}\mathbf{M}_{2D} + \ell_{ex}^2 k\mathbf{L}_{2D}) \in \mathbb{R}^{2N \times 2N}$ to a vector in $\mathbb{R}^{2N}$. Hence, exploiting the block diagonal structure of the preconditioners discussed in Remark 4.5.1, employing $\mathbf{P_Q}[\cdot]$ corresponds to solving a sparse $N \times N$ system with three different right-hand sides, while employing $\mathbf{P}_{2D}$ requires the solution of the same $N \times N$ system for only two right-hand

sides.

Even though the theoretical result for the practical preconditioner (Theorem 4.4.1) is stronger than that for the stationary preconditioner (Theorem 4.4.3), the overall recommendation from the numerical experiments is to use the stationary preconditioner together with an adaptive choice of **T**.

5 Unconditional well-posedness and IMEX improvement of a family of predictor-corrector methods in micromagnetics

This chapter consists of a recently submitted preprint [MPPR21] together with Norbert J. Mauser, Dirk Praetorius, and Michele Ruggeri.

Abstract. Recently, *Kim & Wilkening (Convergence of a mass-lumped finite element method for the Landau–Lifshitz equation, Quart. Appl. Math., 76, 383–405, 2018)* proposed two novel predictor-corrector methods for the Landau–Lifshitz–Gilbert equation (LLG) in micromagnetics, which models the dynamics of the magnetization in ferromagnetic materials. Both integrators are based on the so-called Landau–Lifshitz form of LLG, use mass-lumped variational formulations discretized by first-order finite elements, and only require the solution of linear systems, despite the nonlinearity of LLG. The first(-order in time) method combines a linear update with an explicit projection of an intermediate approximation onto the unit sphere in order to fulfill the LLG-inherent unit-length constraint at the discrete level. In the second(-order in time) integrator, the projection step is replaced by a linear constraint preserving variational formulation. In this paper, we extend the analysis of the integrators by proving unconditional well-posedness and by establishing a close connection of the methods with other approaches available in the literature. Moreover, the new analysis also provides a well-posed integrator for the Schrödinger map equation (which is the limit case of LLG for vanishing damping). Finally, we design an implicit-explicit strategy for the treatment of the lower-order field contributions, which significantly reduces the computational cost of the schemes, while preserving their theoretical properties.

5.1 Introduction

5.1.1 Dynamic micromagnetism

Reliable numerical simulations of magnetic processes occurring at submicrometer length scales are fundamental tools to optimize the design of many technological devices, e.g., magnetic sensors, magnetic logic gates, and hard disk drives. The theoretical background of most simulation packages is the theory of *micromagnetism* [Bro63], a continuum theory which models the magnetic state of a ferromagnetic material at constant temperature in terms of a continuous vector field with constant magnitude, the *magnetization*. A well-accepted model to describe the dynamics of the magnetization is a nonlinear parabolic

partial differential equation (PDE) usually referred to as *Landau–Lifshitz–Gilbert equation (LLG)* [LL35, Gil04], which in the so-called *Landau–Lifshitz (LL) form* reads as

$$\partial_t \boldsymbol{m} = -\frac{1}{1+\alpha^2}\, \boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - \frac{\alpha}{1+\alpha^2}\, \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})). \tag{5.1}$$

Here, $\boldsymbol{m}$ denotes the normalized magnetization, which satisfies the nonconvex unit-length constraint $|\boldsymbol{m}| = 1$, $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m})$ is the effective field, whose specific expression depends on the Gibbs free energy of the system (see (5.6) below), and $\alpha \geq 0$ is the Gilbert damping parameter, which incorporates energy dissipation into the model.

Alternative forms of LLG used in the literature, mathematically equivalent to the LL form (5.1), are the so-called *Gilbert form* of LLG

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) + \alpha\, \boldsymbol{m} \times \partial_t \boldsymbol{m}\,, \tag{5.2}$$

and

$$\alpha\, \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = -\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})), \tag{5.3}$$

which we call the *alternative form* of LLG.

The aforementioned need of fast and reliable tools to perform micromagnetic simulations encouraged many works concerned with the numerical analysis of LLG, which will also be the subject of the present paper.

5.1.2 State of the art

In the last three decades, mathematical questions arising from the micromagnetic theory have been the subject of several studies, from both the analytical and the numerical point of view. For analytical results for LLG, we refer, e.g., to the papers [Vis85, AS92, GH93, CF01a, Mel05, DS14, FT17b, DIP20] and the references therein. For an overview of numerical methods proposed for LLG (up to 2008), we refer to the monograph [Pro01] and the review articles [KP06, GC07, Cim08]. More recently, several numerical schemes with a rigorous convergence analysis have been proposed. They differ from each other in the LLG formulation they are based on (usually one among (5.1)–(5.3)), in the approach used to impose the unit-length constraint at the discrete level, and in the type of convergence result (plain convergence towards a weak solution of LLG with minimal regularity or convergence with rates towards a sufficiently regular strong solution).

Semi-implicit finite element methods based on (variants of) the LL form (5.1) of LLG are proposed in [Gao14, An16], where *a priori* error estimates, which show their convergence towards a smooth solution of LLG, are also established.

A class of methods referred to as *tangent plane schemes* or *projection methods* [AJ06, BKP08, Alo08, BFF+14, AHP+14, AKST14, FT17a, DPP+20, AFKL21] is based on a predictor-corrector approach: At each time-step, first, an update is computed by solving a linear variational problem posed in the discrete tangent space of the current magnetization; second, the update is used to obtain the magnetization at the next time-step. The methods proposed in [AJ06, BKP08, Alo08, BFF+14, AHP+14, FT17a] are based on a variational formulation of (5.3) discretized by first-order finite elements to compute an approximation of the linear velocity $\partial_t \boldsymbol{m}$. The magnetization at the next time-step

is then obtained via a first-order time-stepping. To impose the unit-length constraint at the vertices of the underlying mesh, the nodal values are projected onto the sphere in [AJ06, BKP08, Alo08, BFF$^+$14]. The projection is omitted from the time-stepping in [AHP$^+$14, FT17a]: In this case, the approximations do not fulfill the constraint (not even at the vertices of the mesh), but this error can be controlled by the time-step size (in particular, the constraint holds for the solution of LLG towards which the finite element approximation converges). High-order extensions of the tangent plane approach have been proposed in [AKST14, DPP$^+$20, AFKL21]. The main advantages of this class of methods are that they do not require any time-step restriction for convergence (*unconditional* convergence) [Alo08, BFF$^+$14, AHP$^+$14, AKST14, FT17a, DPP$^+$20] and that, despite the nonlinear nature of LLG, only one linear system per time-step has to be solved.

A numerical scheme based on the Gilbert form (5.2) of LLG is considered in [BP06, PRS18b]. The method employs mass-lumped first-order finite elements for the spatial discretization and the second-order implicit midpoint rule for the time discretization. The scheme is unconditionally convergent towards a weak solution of LLG, but requires the solution of a nonlinear system of equations per time-step. A similar method, but based on the LL form (5.1) of LLG, is proposed and analyzed in [Cim09]. The latter approach is motivated by the interest in having an integrator which is robust with respect to the limit cases of (5.1) in which one of the two terms on the right-hand side tends to zero. Indeed, in the case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \boldsymbol{\Delta m}$, neglecting the second (dissipative) term on the left-hand side of (5.1) ($\alpha \to 0$), one obtains the so-called Schrödinger map equation [SSB86], whereas omitting the first (conservative) term, one is led to the harmonic map heat flow [LW08].

The recent work [KW18] proposes two predictor-corrector schemes for LLG which aim to combine the features of some of the above integrators. In the first scheme, [KW18, Algorithm 1], which we denote by `PC1` for the sake of brevity, the predictor is based on the LL form (5.1) of LLG (like the variational formulation used in [Cim09]) and employs mass-lumping for its discretization (like [BP06, Cim09]). However, it only requires the solution of one linear system per time-step and uses the nodal projection to impose the unit-length constraint (like the method of [Alo08, BFF$^+$14]). The second scheme, [KW18, Algorithm 2], which we refer to as `PC2`, uses the same predictor as `PC1`, but replaces the nodal projection step with a constraint preserving mass-lumped (as in [BP06, Cim09]), but linear (as in [Alo08, BFF$^+$14]), variational formulation. In the paper, adapting the proof of [Alo08], the authors show convergence of the approximations generated by `PC1` towards a weak solution of LLG. Moreover, the expected convergence order in time of both methods (first-order for `PC1`, second-order for `PC2`) is empirically verified by means of numerical experiments in 2D.

Note that in the above discussion we have restricted ourselves to methods employing the finite element method for the spatial discretization. For other approaches based on finite differences, we refer, e.g., to [WGCE01, DSM05, KL17, XGCW$^+$20, CWX21] and the references therein.

5.1.3 Novelty of the present work

In this work, we improve the theoretical understanding of the predictor-corrector methods proposed in [KW18].

First, we show that `PC1` is unconditionally well-posed, i.e., for each time-step, the variational problem to be solved admits a unique solution, which is left open in the original paper. By closing this fundamental gap, we show that `PC1` is not only closely related to the first-order tangent plane scheme of [Alo08, BFF$^+$14], but actually can even be interpreted as a slight modification of it, which explains why the convergence analysis of the two schemes is almost identical. Furthermore, following [BFF$^+$14], we propose an implicit-explicit (IMEX) version of `PC1`. When considering magnetization dynamics involving the full effective field—more precisely, dynamics including the nonlocal stray field—the proposed adaptation is computationally much more attractive: The IMEX version `PC1+IMEX` avoids the costly inner iteration in the solver of the original scheme, while preserving the experimental first-order accuracy of `PC1`, which we confirm by numerical studies in 3D.

Second, we consider the analysis of `PC2`. While the conservation of the unit-length constraint at the vertices of the mesh in `PC1` is guaranteed (at machine precision) also in practical computations (since it is directly enforced in the method using the nodal projection), the one guaranteed by `PC2`, which follows from the variational formulation of the corrector, is lost in practice due to the inevitable use of inexact (iterative) solvers for the solution of the arising linear systems. Hence, although the predictors of `PC1` and `PC2` coincide in theory, the well-posedness analysis of (the predictor of) `PC1` does not transfer to a practical version of `PC2`. To cope with this problem, we establish a decomposition of the finite element space, which does not only allow us to prove unconditional well-posedness of the practical version of `PC2`, but also to extend the result, for both `PC1` and `PC2` (theoretical and practical), to the limit case $\alpha = 0$ (Schrödinger map equation). Moreover, following [PRS18b, DPP$^+$20], we adopt the IMEX treatment also for `PC2`. In particular, in the presence of the nonlocal stray field, the proposed method `PC2+IMEX` is computationally much more attractive than its fully implicit counterpart `PC2`, while conserving the experimental second-order accuracy in time. Again, these claims are confirmed in our numerical studies. Stability and convergence of `PC2`, not addressed in [KW18], remain open also in our analysis and will be the subject of future research. In this paper, we shed some light on this question by means of some surprising numerical experiments.

5.1.4 Outline

We conclude this section by collecting some general notation and basic vector identities used throughout the work (Section 5.1.5). In Section 5.2, we formulate the initial boundary value problem for LLG in which we are interested, we recall the notion of a weak solution and introduce the basic ingredients of the discretization. Section 5.3 is devoted to the first-order method: After proving unconditional well-posedness of `PC1` in Section 5.3.2, we propose an IMEX adaptation (Section 5.3.3) overcoming the inefficiency drawbacks of the original version, while preserving unconditional well-posedness, stability, convergence (Section 5.3.4), and accuracy. Section 5.4 is devoted to the second-order method: In Section 5.4.2, we first prove unconditional well-posedness of `PC2`. Subsequently, in Section 5.4.3, we extend the unconditional well-posedness result to the more general formulation of the second-order algorithm, where discrete unit-length of the iterates is not assumed. This covers, in particular, the practical version of the scheme incorporating the inevitable use of inexact (iterative) linear solvers. Section 5.4.4 closes with a second-order accuracy preserv-

ing IMEX modification overcoming the inefficiency drawbacks of `PC2`. Section 5.5 provides numerical studies validating the applicability (Section 5.5.1) and the expected accuracy (Section 5.5.2) of the IMEX integrators proposed in this work. Finally, in Section 5.5.3, we numerically investigate the stability of `PC2`.

5.1.5 General notation and vector identities

Throughout this work, we use the standard notation for Lebesgue, Sobolev, and Bochner spaces and norms. Vector-valued functions are indicated by bold letters. Bold letters are also used for vector-valued and matrix-valued function spaces, e.g., both $L^2(\Omega; \mathbb{R}^3)$ and $L^2(\Omega; \mathbb{R}^{3\times3})$ are denoted by $\boldsymbol{L}^2(\Omega)$. We denote by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ the scalar product and the norm of $\boldsymbol{L}^2(\Omega)$, respectively, while $|\cdot|$ denotes the Euclidean norm of a vector in $\mathbb{R}^3$ or the Frobenius norm of a matrix in $\mathbb{R}^{3\times3}$. To abbreviate notation in proofs, we write $A \lesssim B$ when $A \leq cB$ for some generic constant $c > 0$, which is clear from the context and always independent of the discretization parameters. For vector-valued functions $\boldsymbol{f}, \boldsymbol{g} \colon \Omega \to \mathbb{R}^3$ we use the notation

$$-\boldsymbol{g} \times \boldsymbol{\nabla}\boldsymbol{f} := \boldsymbol{\nabla}\boldsymbol{f} \times \boldsymbol{g} := (\partial_1\boldsymbol{f} \times \boldsymbol{g}, \partial_2\boldsymbol{f} \times \boldsymbol{g}, \partial_3\boldsymbol{f} \times \boldsymbol{g}) \colon \Omega \to \mathbb{R}^{3\times3}.$$

We conclude this section by recalling five vector identities used regularly in this work

$$\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}, \tag{5.4a}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{a} = 0, \tag{5.4b}$$

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\,\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\,\boldsymbol{c}, \tag{5.4c}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}), \tag{5.4d}$$

$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) = (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{b} \cdot \boldsymbol{c})(\boldsymbol{a} \cdot \boldsymbol{d}), \tag{5.4e}$$

which hold true for arbitrary $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d} \in \mathbb{R}^3$.

5.2 Problem formulation

5.2.1 Landau–Lifshitz–Gilbert equation

Given a bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$ and $T > 0$, we define the space-time cylinder $\Omega_T := \Omega \times (0, T)$. We consider the following initial boundary value problem

$$(1 + \alpha^2)\,\partial_t\boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - \alpha\,\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})) \qquad \text{in } \Omega_T, \tag{5.5a}$$

$$\partial_{\boldsymbol{n}}\boldsymbol{m} = \boldsymbol{0} \qquad \text{on } \partial\Omega \times (0, T), \tag{5.5b}$$

$$\boldsymbol{m}(0) = \boldsymbol{m}^0 \qquad \text{in } \Omega. \tag{5.5c}$$

The unknown is the normalized magnetization $\boldsymbol{m} \colon \Omega_T \to \mathbb{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3 \colon |\boldsymbol{x}| = 1\}$. In (5.5a), the effective field

$$\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2\,\boldsymbol{\Delta}\boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f} \tag{5.5d}$$

is the negative functional derivative of the Gibbs free energy

$$\mathcal{E}(\boldsymbol{m}) = \frac{\ell_{\text{ex}}^2}{2} \int_\Omega |\boldsymbol{\nabla}\boldsymbol{m}|^2\,dx - \frac{1}{2}\int_\Omega \boldsymbol{\pi}(\boldsymbol{m}) \cdot \boldsymbol{m}\,dx - \int_\Omega \boldsymbol{f} \cdot \boldsymbol{m}\,dx, \tag{5.6}$$

where $\ell_{\mathrm{ex}} > 0$ is the exchange length, $\boldsymbol{\pi} \colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ is a linear, continuous, and self-adjoint operator which collects all lower-order contributions such as uniaxial magnetocrystalline anisotropy and the nonlocal stray field, and $\boldsymbol{f} \colon \Omega_T \to \mathbb{R}^3$ is the applied external field. The equation is supplemented with homogeneous Neumann boundary conditions (5.5b) and the initial condition (5.5c), where $\boldsymbol{m}^0 \colon \Omega \to \mathbb{S}^2$ denotes a given initial state.

Taking the scalar product of (5.5a) with $\boldsymbol{m}$, (5.4b) yields that $0 = \partial_t \boldsymbol{m} \cdot \boldsymbol{m}$ in Ω_T. Since $|\boldsymbol{m}^0| = 1$ in Ω by assumption and $\partial_t(|\boldsymbol{m}|^2/2) = \partial_t \boldsymbol{m} \cdot \boldsymbol{m} = 0$, it follows that $|\boldsymbol{m}| = 1$ in Ω_T. Moreover, any solution of (5.5a) satisfies the energy law

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(\boldsymbol{m}, \boldsymbol{f}) = -\alpha \int_\Omega |\partial_t \boldsymbol{m}|^2 \, dx - \int_\Omega \partial_t \boldsymbol{f} \cdot \boldsymbol{m} \, dx. \tag{5.7}$$

From this, we see that the Gilbert damping constant α modulates the dissipation of the system. In particular, if $\alpha = 0$ and $\boldsymbol{f}$ is constant in time, then the energy is conserved. The PDE inherent constraint $|\boldsymbol{m}| = 1$ in Ω_T and the energy law (5.7) should be satisfied (at the discrete level) by any feasible numerical method.

5.2.2 Weak solution

We recall the notion of a weak solution of (5.5), which extends the one introduced in [AS92].

Definition 5.2.1. *Let $\boldsymbol{m}^0 \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ and $\boldsymbol{f} \in C^1([0, T]; \boldsymbol{L}^2(\Omega))$. A vector field $\boldsymbol{m} \colon \Omega_T \to \mathbb{R}$ is called a weak solution of (5.5), if the following properties are satisfied:*

(i) $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T) \cap L^\infty(0, T; \boldsymbol{H}^1(\Omega))$ *with* $|\boldsymbol{m}| = 1$ *a.e. in* Ω_T;

(ii) $\boldsymbol{m}(0) = \boldsymbol{m}^0$ *in the sense of traces;*

(iii) for all $\boldsymbol{w} \in \boldsymbol{H}^1(\Omega_T)$, *it holds that*

$$\int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{w}(t) \rangle \, \mathrm{d}t - \alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{w}(t) \rangle \, \mathrm{d}t$$
$$= \ell_{\mathrm{ex}}^2 \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{\nabla} \boldsymbol{m}(t), \boldsymbol{\nabla} \boldsymbol{w}(t) \rangle \, \mathrm{d}t - \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{\pi}(\boldsymbol{m}(t)), \boldsymbol{w}(t) \rangle \, \mathrm{d}t \tag{5.8}$$
$$- \int_0^T \langle \boldsymbol{m}(t) \times \boldsymbol{f}(t), \boldsymbol{w}(t) \rangle \, \mathrm{d}t;$$

(iv) it holds that

$$\mathcal{E}(\boldsymbol{m}(T)) + \alpha \int_0^T \|\partial_t \boldsymbol{m}(t)\|^2 \mathrm{d}t + \int_0^T \langle \partial_t \boldsymbol{f}(t), \boldsymbol{m}(t) \rangle \, \mathrm{d}t \leq \mathcal{E}(\boldsymbol{m}^0). \tag{5.9}$$

We note that (5.8) is a variational formulation in space-time of the Gilbert form (5.2) of LLG, and that (5.9) is a weaker version of the energy law (5.7).

5.2.3 Discretization

For the temporal discretization, given $L \in \mathbb{N}$, we consider a partition $\{t_\ell\}_{\ell=0,\dots,L}$ of the time interval $[0, T]$ with uniform time-step size $k := T/L > 0$, i.e., $t_\ell = \ell k$ for all $\ell = 0, \dots, L$. Given a finite sequence of functions $\{\boldsymbol{u}^\ell\}_{\ell=0,\dots,L}$, we define

$$\boldsymbol{u}^{\ell+1/2} := \frac{\boldsymbol{u}^{\ell+1} + \boldsymbol{u}^\ell}{2} \quad \text{and} \quad d_t \boldsymbol{u}^{\ell+1} := \frac{\boldsymbol{u}^{\ell+1} - \boldsymbol{u}^\ell}{k} \quad \text{for all } \ell = 0, \dots L - 1.$$

For the spatial discretization, we consider a regular tetrahedral triangulation $\mathcal{T}_h$ of Ω with mesh size $h > 0$. We denote by $\mathcal{N}_h$ the set of vertices of $\mathcal{T}_h$ and by $\{\phi_{\boldsymbol{z}}\}_{\boldsymbol{z} \in \mathcal{N}_h}$ the classical nodal basis of the space $\mathcal{S}^1(\mathcal{T}_h)$ of $\mathcal{T}_h$-piecewise linear and globally continuous discrete functions, i.e., $\phi_{\boldsymbol{z}}(\boldsymbol{z}') = \delta_{\boldsymbol{z},\boldsymbol{z}'}$ for all $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{N}_h$. With $\{\boldsymbol{e}_j\}_{j=1,2,3}$ the standard basis of $\mathbb{R}^3$, $\{\phi_{\boldsymbol{z}} \boldsymbol{e}_j\}_{\boldsymbol{z} \in \mathcal{N}_h, j=1,2,3}$ gives a basis of $\mathcal{S}^1(\mathcal{T}_h)^3$. Note that $\mathcal{S}^1(\mathcal{T}_h)^3$ is a $3N$-dimensional space, with N denoting the number of vertices in $\mathcal{N}_h$. We introduce the *set of admissible discrete magnetizations*

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

and, for $\boldsymbol{m}_h \in \mathcal{M}_h$, the *discrete tangent space* of $\boldsymbol{m}_h$

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

We consider the nodal interpolant $\mathcal{I}_h \colon C^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)$, which is defined by $\mathcal{I}_h(v) = \sum_{\boldsymbol{z} \in \mathcal{N}_h} v(\boldsymbol{z})\phi_{\boldsymbol{z}}$ for all $v \in C^0(\overline{\Omega})$. We denote the vector-valued realization of the nodal interpolant by $\boldsymbol{\mathcal{I}}_h \colon \boldsymbol{C}^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)^3$. In $\boldsymbol{C}^0(\overline{\Omega})$, besides the standard $\boldsymbol{L}^2(\Omega)$-scalar product $\langle \cdot, \cdot \rangle$, we consider the mass-lumped scalar product $\langle \cdot, \cdot \rangle_h$ defined by

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle_h = \int_\Omega \boldsymbol{\mathcal{I}}_h(\boldsymbol{u} \cdot \boldsymbol{w}) \, dx \quad \text{for all } \boldsymbol{u}, \boldsymbol{w} \in \boldsymbol{C}^0(\overline{\Omega}).$$

Using the definition of the nodal interpolant, we see that

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle_h = \sum_{\boldsymbol{z} \in \mathcal{N}_h} \beta_{\boldsymbol{z}} \, \boldsymbol{u}(\boldsymbol{z}) \cdot \boldsymbol{w}(\boldsymbol{z}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{w} \in \boldsymbol{C}^0(\overline{\Omega}), \tag{5.10}$$

where $\beta_{\boldsymbol{z}} := \int_\Omega \phi_{\boldsymbol{z}} \, dx > 0$ for all $\boldsymbol{z} \in \mathcal{N}_h$. For discrete functions, the induced norm $\|\cdot\|_h$ is equivalent to the standard $\boldsymbol{L}^2(\Omega)$-norm; see [Bar15, Lemma 3.9], i.e., it holds that

$$\|\boldsymbol{w}_h\| \leq \|\boldsymbol{w}_h\|_h \leq \sqrt{5}\|\boldsymbol{w}_h\| \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{5.11}$$

We define the (negative) discrete Laplacian $-\boldsymbol{\Delta}_h \colon \boldsymbol{H}^1(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)^3$ by

$$-\langle \boldsymbol{\Delta}_h \boldsymbol{w}, \boldsymbol{w}_h \rangle_h = \langle \boldsymbol{\nabla} \boldsymbol{w}, \boldsymbol{\nabla} \boldsymbol{w}_h \rangle \quad \text{for all } \boldsymbol{w} \in \boldsymbol{H}^1(\Omega) \text{ and } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{5.12}$$

Let $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$. With a double application of the classical inverse estimate and the norm equivalence (5.11), we see that

$$\|\boldsymbol{\Delta}_h \boldsymbol{w}_h\|_h^2 = \langle \boldsymbol{\Delta}_h \boldsymbol{w}_h, \boldsymbol{\Delta}_h \boldsymbol{w}_h \rangle_h \overset{(5.12)}{=} -\langle \boldsymbol{\nabla} \boldsymbol{w}_h, \boldsymbol{\nabla} \boldsymbol{\Delta}_h \boldsymbol{w}_h \rangle \leq \|\boldsymbol{\nabla} \boldsymbol{w}_h\| \|\boldsymbol{\nabla} \boldsymbol{\Delta}_h \boldsymbol{w}_h\|$$
$$\leq C h^{-2} \|\boldsymbol{w}_h\|_h \|\boldsymbol{\Delta}_h \boldsymbol{w}_h\|_h.$$

This shows that

$$\|\boldsymbol{\Delta}_h \boldsymbol{w}_h\|_h \le Ch^{-2}\|\boldsymbol{w}_h\|_h \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3, \tag{5.13}$$

where $C > 0$ depends only on the quasi-uniformity of the triangulation $\mathcal{T}_h$. Finally, we define the mapping $\mathbb{P}_h \colon \boldsymbol{L}^2(\Omega) \to \mathcal{S}^1(\mathcal{T}_h)^3$ by

$$\langle \mathbb{P}_h \boldsymbol{w}, \boldsymbol{w}_h \rangle_h = \langle \boldsymbol{w}, \boldsymbol{w}_h \rangle \quad \text{for all } \boldsymbol{w} \in \boldsymbol{L}^2(\Omega) \text{ and } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{5.14}$$

Using (5.10), it is easy to see that, for all $\boldsymbol{w} \in \boldsymbol{L}^2(\Omega)$ and all $\boldsymbol{z} \in \mathcal{N}_h$, it holds that $(\mathbb{P}_h \boldsymbol{w})(\boldsymbol{z}) = \beta_{\boldsymbol{z}}^{-1} \int_\Omega \boldsymbol{w} \phi_{\boldsymbol{z}} \, dx$. In particular, the computation of $\mathbb{P}_h \boldsymbol{w}$ does not require to solve any linear system.

5.3 First-order predictor-corrector scheme

In this section, we discuss the first-order scheme proposed in [KW18] and its connections with the integrators proposed in [BP06] and [Alo08]. Our contribution is twofold: First, we prove unconditional well-posedness of the scheme, which fills a fundamental gap in the analysis of [KW18]. Second, we employ an explicit treatment of the (nonlocal) lower-order contributions to obtain a computationally superior IMEX version of the scheme, preserving (unconditional) convergence and experimental rates in time. We first consider the method for the case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \boldsymbol{\Delta m}$. For the general case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \boldsymbol{\Delta m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$, we refer to Section 5.3.3.

5.3.1 Variational formulation

The following algorithm restates [KW18, Algorithm 1] written in terms of the discrete functions $\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$, where $\boldsymbol{m}_h^\ell \approx \boldsymbol{m}(t_\ell)$, $\boldsymbol{v}_h^\ell \approx \partial_t \boldsymbol{m}(t_\ell)$, and $\boldsymbol{m}_h^{\ell+1} \approx \boldsymbol{m}(t_{\ell+1})$. In particular, the predictor (5.15) of Algorithm 5.3.1 reformulates the N equations in $\mathbb{R}^3$ of the predictor of [KW18, Algorithm 1] as an equivalent variational formulation for $\boldsymbol{v}_h^\ell$ in $\mathcal{S}^1(\mathcal{T}_h)^3$. As for the tangent plane scheme [Alo08], $\theta \in [0,1]$ is a parameter modulating the 'degree of implicitness' of the scheme.

Algorithm 5.3.1 (PC1, variational form). *__Input:__ $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.*
__Loop:__ For all time-steps $\ell = 0, \dots, L-1$, iterate:

(i) *Compute $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that*

$$\begin{aligned}
(1+\alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h = &-\ell_{\text{ex}}^2 \langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell), \boldsymbol{w}_h \rangle_h \\
&- \alpha \ell_{\text{ex}}^2 \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell)), \boldsymbol{w}_h \rangle_h.
\end{aligned} \tag{5.15}$$

(ii) *Define $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ by*

$$\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})}{|\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h. \tag{5.16}$$

__Output:__ Sequence of discrete functions $\left\{ (\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1}) \right\}_{\ell=0}^{L-1}$.

5.3.2 Unconditional well-posedness

The predictor (5.15) can be written as: Find $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{w}_h) = F_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}_h) \qquad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3\,,$$

with some linear form $F_{\mathrm{pre}}[\boldsymbol{m}_h^\ell]$ and the bilinear form $a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell]$ on $\mathcal{S}^1(\mathcal{T}_h)^3$ reading

$$\begin{aligned}
a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{w}_h) := {}& (1 + \alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h + \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h \\
& + \alpha \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{v}_h^\ell), \boldsymbol{w}_h \rangle_h\,.
\end{aligned}$$

From the boundedness of $\boldsymbol{m}_h^\ell$ in $\boldsymbol{L}^\infty(\Omega)$ guaranteed by the nodal projection (5.16) and an inverse estimate on the discrete Laplacian (5.13) we have

$$a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}_h, \boldsymbol{w}_h) \geq (1 - Ckh^{-2})\|\boldsymbol{w}_h\|_h^2\,.$$

Hence, assuming the CFL condition $k = o(h^2)$ implies the coercivity of $a_{\mathrm{pre}}[\boldsymbol{m}_h^\ell]$ for sufficiently small h and k. However, this undesirable restriction is a consequence of naively using the inverse estimate, and can be avoided.

For arbitrary $\alpha > 0$ the upcoming refined analysis allows to drop any CFL-type assumptions on the discretization parameters: In Lemma 5.3.2, we first collect two basic properties of Algorithm 5.3.1, which turn out to be sufficient to prove unconditional well-posedness of the algorithm in Theorem 5.3.3; also see Remark 5.3.4.

Lemma 5.3.2. *Let $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$. Suppose that the solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ to (5.15) exists. Then, $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, and (5.16) provides a well-defined $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.*

Proof. For arbitrary $\boldsymbol{z} \in \mathcal{N}_h$, with $\phi_{\boldsymbol{z}} \in \mathcal{S}^1(\mathcal{T}_h)$ denoting the hat function with $\phi_{\boldsymbol{z}}(\boldsymbol{z}') = \delta_{\boldsymbol{z},\boldsymbol{z}'}$ for all $\boldsymbol{z}' \in \mathcal{N}_h$, we choose $\boldsymbol{w}_h := \boldsymbol{m}_h^\ell(\boldsymbol{z})\phi_{\boldsymbol{z}} \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (5.15) to see

$$\boldsymbol{m}_h^\ell(\boldsymbol{z}) \cdot \boldsymbol{v}_h^\ell(\boldsymbol{z}) \stackrel{(5.10)}{=} \beta_{\boldsymbol{z}}^{-1}\langle \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell(\boldsymbol{z})\phi_{\boldsymbol{z}} \rangle_h \stackrel{(5.15),(5.4\mathrm{b})}{=} 0\,.$$

Hence, $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ belongs to $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$.
Well-posedness of (5.16) follows immediately from $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ via

$$|\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k\boldsymbol{v}_h^\ell(\boldsymbol{z})|^2 = |\boldsymbol{m}_h^\ell(\boldsymbol{z})|^2 + k^2|\boldsymbol{v}_h^\ell(\boldsymbol{z})|^2 \geq |\boldsymbol{m}_h^\ell(\boldsymbol{z})|^2 = 1 \quad \text{for all } \ell = 0, \ldots, L-1\,.$$

Consequently, for all $\boldsymbol{z} \in \mathcal{N}_h$ the denominator in (5.16) is bounded below by $|\boldsymbol{m}_h^\ell(\boldsymbol{z})| = 1$ and the corrector step of Algorithm 5.3.1 is always well-posed.
The third claim $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ follows directly from the explicit projection in (5.16). $\qquad\square$

These two observations are already sufficient to prove the first main contribution of this work.

Theorem 5.3.3. *Let $\alpha > 0$. Then, Algorithm 5.3.1 is unconditionally well-posed for any input $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, i.e., for all $\ell = 0, \ldots, L-1$ the predictor (5.15) admits a unique solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ and the corrector (5.16) is well-posed providing $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.*

Proof. Well-posedness of the corrector (5.16) and $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ follow from Lemma 5.3.2. Transforming (5.15) into a coercive system in the discrete tangent space, we prove well-posedness of the predictor in three steps:

- **Step 1:** The predictor of Algorithm 5.3.1 can be reformulated as a well-posed system.

We claim that $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ satisfies (5.15) for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, if and only if it satisfies $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ as well as

$$\alpha\langle\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h + \langle\boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h = \ell_{\mathrm{ex}}^2\langle\boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h\rangle_h \quad \text{for all } \boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]. \quad (5.17)$$

This formulation can be written as follows: Find $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ such that

$$a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h) = \ell_{\mathrm{ex}}^2\langle\boldsymbol{\Delta}_h\boldsymbol{m}_h^\ell, \boldsymbol{\varphi}_h\rangle_h \quad \text{for all } \boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell],$$

where the bilinear form $a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell]\colon \mathcal{K}_h[\boldsymbol{m}_h^\ell] \times \mathcal{K}_h[\boldsymbol{m}_h^\ell] \to \mathbb{R}$ is defined by

$$a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell](\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h) := \alpha\langle\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h + \langle\boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h - \ell_{\mathrm{ex}}^2\theta k\langle\boldsymbol{\Delta}_h\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h.$$

For $\alpha > 0$, the bilinear form satisfies the ellipticity property

$$a_{\mathrm{alt}}[\boldsymbol{m}_h^\ell](\boldsymbol{\varphi}_h, \boldsymbol{\varphi}_h) = \alpha\|\boldsymbol{\varphi}_h\|_h^2 + \ell_{\mathrm{ex}}^2\theta k\|\boldsymbol{\nabla}\boldsymbol{\varphi}_h\|^2 \quad \text{for all } \boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell],$$

and the problem (5.17) is well-posed by the Lax–Milgram theorem. To conclude the proof, it remains to show the claimed equivalence of (5.15) and (5.17).

- **Step 2:** Any solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ of (5.15) also solves (5.17).

Given arbitrary $\boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, we choose $\boldsymbol{w}_h = \mathcal{I}_h(\alpha\boldsymbol{\varphi}_h + \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell) \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (5.15) to obtain

$$\begin{aligned}
(1 + \alpha^2)\alpha\langle\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h + (1 + \alpha^2)&\langle\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell\rangle_h = -\alpha\ell_{\mathrm{ex}}^2\langle\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h\rangle_h \\
&- \ell_{\mathrm{ex}}^2\langle\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell\rangle_h - \alpha^2\ell_{\mathrm{ex}}^2\langle\boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell)), \boldsymbol{\varphi}_h\rangle_h \\
&- \alpha\ell_{\mathrm{ex}}^2\langle\boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell)), \boldsymbol{\varphi}_h \times \boldsymbol{m}_h^\ell\rangle_h.
\end{aligned} \quad (5.18)$$

By (5.4d) the left-hand side of (5.18) resembles the left-hand side of (5.17) scaled by $(1+\alpha^2)$. From $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{\varphi}_h \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, we infer $\mathcal{I}_h(|\boldsymbol{m}_h^\ell|^2) = 1$ and $\mathcal{I}_h(\boldsymbol{m}_h^\ell \cdot \boldsymbol{\varphi}_h) = 0$ in Ω. Hence, using the vector identities (5.4b)–(5.4e), the first and the last term on the right-hand side of (5.18) cancel out, and (5.18) equivalently reads

$$(1 + \alpha^2)\big(\alpha\langle\boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h + \langle\boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{\varphi}_h\rangle_h\big) = (1 + \alpha^2)\ell_{\mathrm{ex}}^2\langle\boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell), \boldsymbol{\varphi}_h\rangle_h.$$

Now multiplying (5.18) by $1/(1 + \alpha^2)$, we conclude that any $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ satisfying (5.15) necessarily satisfies (5.17) and, according to Lemma 5.3.2, belongs to $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$ itself.

- **Step 3:** Any solution $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ of (5.17) also solves (5.15).

Given arbitrary $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, we choose $\boldsymbol{\varphi}_h = \mathcal{I}_h\big(\boldsymbol{m}_h^\ell \times \boldsymbol{w}_h + \alpha \boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell)\big) \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ in (5.17) to obtain

$$
\begin{aligned}
\alpha \langle \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times \boldsymbol{w}_h \rangle_h &+ \alpha^2 \langle \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell) \rangle_h \\
&+ \langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times \boldsymbol{w}_h \rangle_h + \alpha \langle \boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell) \rangle_h \\
&= \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell), \boldsymbol{m}_h^\ell \times \boldsymbol{w}_h \rangle_h + \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell), \boldsymbol{m}_h^\ell \times (\boldsymbol{w}_h \times \boldsymbol{m}_h^\ell) \rangle_h \,.
\end{aligned}
\tag{5.19}
$$

From $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, we infer $\mathcal{I}_h(|\boldsymbol{m}_h^\ell|^2) = 1$ and $\mathcal{I}_h(\boldsymbol{m}_h^\ell \cdot \boldsymbol{v}_h^\ell) = 0$ in Ω. Hence, by the vector identities (5.4b)–(5.4e), the first and the last term on the left-hand side of (5.19) cancel out, while the second and third term on the left-hand side of (5.19) add up to the left-hand side of (5.15). Further, by (5.4d) the right-hand side of (5.19) resembles the right-hand side of (5.15). We conclude that any $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell] \subset \mathcal{S}^1(\mathcal{T}_h)^3$ satisfying (5.17) necessarily satisfies (5.15). Ultimately, we have shown that (5.15) is equivalent to (5.17), which always allows for a unique solution as shown in Step 1. □

Remark 5.3.4. *(i) Let $\boldsymbol{w} \colon \Omega \to \mathbb{R}^3$ be an arbitrary smooth test function. Writing $\boldsymbol{m}^\ell :=\boldsymbol{m}(t_\ell)$ and $\boldsymbol{v}^\ell := \partial_t \boldsymbol{m}(t_\ell)$, the variational formulation of the LL form (5.5a) of LLG at time $t_\ell \in (0, T)$ reads*

$$
(1 + \alpha^2)\langle \boldsymbol{v}^\ell, \boldsymbol{w} \rangle \;\; = \;\; -\ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}^\ell \times \boldsymbol{\Delta m}^\ell, \boldsymbol{w} \rangle - \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}^\ell \times (\boldsymbol{m}^\ell \times \boldsymbol{\Delta m}^\ell), \boldsymbol{w} \rangle \,.
$$

The discrete variational formulation (5.15) can be seen as a discrete mass-lumped version of the latter, where the effective field is treated implicitly in time.

(ii) The core of the proof of Theorem 5.3.3 is the equivalent reformulation of the predictor step (5.15) as well-posed system (5.17) in the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$. For $\alpha > 0$, the reformulated system is unconditionally well-posed and corresponds to a discretization of the alternative form of LLG (5.3). Using (5.4c) and $|\boldsymbol{m}|^2 \equiv 1$, the formulation (5.3) is directly obtained from the LL form (5.5a) via $(\alpha \cdot (5.5a) + \boldsymbol{m} \times (5.5a))/(1 + \alpha^2)$. Step 3 of the proof of Theorem 5.3.3 resembles the analogous computations on a discrete level. We emphasize that the mass-lumped scalar product $\langle \cdot, \cdot \rangle_h$ as well as $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ are the crucial ingredients in the proof of Theorem 5.3.3.

(iii) With the reformulation (5.17), we fully understand the real nature of the first-order integrator from [KW18]: It is a predictor-corrector scheme which combines the approaches of Bartels & Prohl [BP06] (mass-lumping (5.10), discrete Laplacian (5.12)) and Alouges [Alo08] (degree of implicitness θ, projection update (5.16), unknown approximates time derivative). The predictor step (5.15) is a mass-lumped discrete variational formulation of the LL form (5.5a) of LLG. The equivalent variational formulation (5.17) is a mass-lumped variational formulation of the alternative form (5.3) of LLG and, in particular, is the mass-lumped version of the predictor step of the tangent plane scheme from [Alo08]. Analogously to the tangent plane scheme, the corrector step of Algorithm 5.3.1 employs the nodal projection to enforce the modulus constraint at the vertices of the triangulations.

(iv) While the proof of Theorem 5.3.3 emphasizes the close relation of Algorithm 5.3.1 to the first-order tangent plane scheme, it is restricted to $\alpha > 0$. In fact, Theorem 5.3.3 can also be proved for the limit case $\alpha = 0$; see Remark 5.4.6(iii)–(iv) below.

5.3.3 Including lower-order contributions

In this section, we discuss the extension of the scheme to the general case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \boldsymbol{\Delta m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$. We start by recalling the definition (5.14) of the mapping $\mathbb{P}_h \colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ and assume that we are given an operator $\boldsymbol{\pi}_h \colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ which approximates $\boldsymbol{\pi}$, e.g., in the case of the nonlocal stray field $\boldsymbol{\pi}(\boldsymbol{m}) = \boldsymbol{h}_{\text{s}}$, $\boldsymbol{\pi}_h$ is a method for the approximation of the magnetostatic Maxwell equations, e.g., via the hybrid FEM-BEM method from [FK90].

In the original first-order integrator from [KW18], the lower-order contributions are treated implicitly in time. Rewritten as a mass-lumped discrete LL formulation like (5.15), the predictor step of [KW18, Algorithm 1] reads as follows: Find $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$(1 + \alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h = -\langle \boldsymbol{m}_h^\ell \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+\theta})], \boldsymbol{w}_h \rangle_h$$
$$- \alpha\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+\theta})]), \boldsymbol{w}_h \rangle_h \quad (5.20)$$

for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$. Here, $\boldsymbol{f}^{\ell+\theta} = \boldsymbol{f}(t_\ell + \theta k)$ for all $\ell = 0, \ldots, L-1$. However, this approach for the inclusion of the lower-order terms is not very attractive from the computational point of view: Indeed, the variational formulation comprises the term $\boldsymbol{\pi}_h(\boldsymbol{v}_h^\ell)$ which requires to solve a (possibly nonlocal) problem for the unknown. An implementation of this scheme would then be based on a costly inner iteration.

From our previous work on the tangent plane scheme [BFF+14, DPP+20] and on the midpoint scheme [PRS18b], we know that an explicit treatment is favorable: Therefore, we change the above variational formulation: Find $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$(1 + \alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h = -\langle \boldsymbol{m}_h^\ell \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell)], \boldsymbol{w}_h \rangle_h$$
$$- \alpha\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell)]), \boldsymbol{w}_h \rangle_h$$

for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$. Only the leading-order exchange contribution is treated implicitly in time, while the lower-order contributions are treated explicitly. This does not spoil the convergence result of the scheme (since the nodal projection already restricts the scheme to first-order in time) and it is computationally much more attractive. To sum up, we consider the following implicit-explicit (IMEX) algorithm.

Algorithm 5.3.5 (PC1+IMEX). ***Input:*** $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: *For all time-steps $\ell = 0, \ldots, L-1$, iterate:*

(i) *Compute $\mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell)) \in \mathcal{S}^1(\mathcal{T}_h)^3$.*

(ii) *Compute $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that*
$$(1 + \alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h = -\langle \boldsymbol{m}_h^\ell \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell)], \boldsymbol{w}_h \rangle_h \quad (5.21)$$
$$- \alpha\langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\text{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell)]), \boldsymbol{w}_h \rangle_h \, .$$

(iii) *Define $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ by*
$$\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k \boldsymbol{v}_h^\ell(\boldsymbol{z})}{|\boldsymbol{m}_h^\ell(\boldsymbol{z}) + k \boldsymbol{v}_h^\ell(\boldsymbol{z})|} \in \mathbb{S}^2 \quad \text{for all } \boldsymbol{z} \in \mathcal{N}_h. \quad (5.22)$$

Output: *Sequence of discrete functions $\left\{ (\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1}) \right\}_{\ell=0}^{L-1}$.*

5.3.4 Stability of Algorithm 5.3.5

Well-posedness of Algorithm 5.3.5 follows from well-posedness of Algorithm 5.3.1 (Theorem 5.3.3), as the system matrices for the linear systems corresponding to the left-hand sides of (5.21) and (5.15), respectively, coincide.

For stability of Algorithm 5.3.5, we assume that all off-diagonal entries of the stiffness matrix $A = (a_{\boldsymbol{z},\boldsymbol{z}'})_{\boldsymbol{z},\boldsymbol{z}'\in\mathcal{N}_h}$ are nonpositive, i.e., it holds that

$$a_{\boldsymbol{z},\boldsymbol{z}'} = \langle\nabla\phi_{\boldsymbol{z}'},\nabla\phi_{\boldsymbol{z}}\rangle \leq 0 \quad \text{for all } \boldsymbol{z},\boldsymbol{z}' \in \mathcal{N}_h \text{ with } \boldsymbol{z} \neq \boldsymbol{z}'. \tag{5.23}$$

This requirement, usually referred to as *angle condition*[1], ensures that the nodal projection $\boldsymbol{w}_h \mapsto \boldsymbol{\mathcal{I}}_h[\boldsymbol{w}_h/|\boldsymbol{w}_h|]$ does not increase the exchange energy of a discrete function, i.e., it holds that

$$\|\boldsymbol{\nabla}\boldsymbol{\mathcal{I}}_h[\boldsymbol{w}_h/|\boldsymbol{w}_h|]\| \leq \|\boldsymbol{\nabla}\boldsymbol{w}_h\|, \tag{5.24}$$

for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$ with $|\boldsymbol{w}_h(\boldsymbol{z})| \geq 1$ for all $\boldsymbol{z} \in \mathcal{N}_h$; see [Bar05, Lemma 3.2]. Moreover, we assume that the discrete operator $\boldsymbol{\pi}_h \colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \boldsymbol{L}^2(\Omega)$ is stable in the sense that

$$\|\boldsymbol{\pi}_h(\boldsymbol{w}_h)\| \leq C\|\boldsymbol{w}_h\| \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3, \tag{5.25}$$

which is met in many practical situations; see [BFF$^+$14]. Under these assumptions, there holds stability of Algorithm 5.3.5.

Theorem 5.3.6. *Let $\mathcal{T}_h$ such that (5.24) holds true. For input $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$, let the approximations $\big\{(\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1})\big\}_{\ell=0}^{L-1}$ be the output of Algorithm 5.3.5. Then, for all $J = 0,\ldots,L-1$, there holds the stability estimate*

$$\begin{aligned}
&\frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^J\|^2 + \alpha k \sum_{\ell=0}^{J-1}\|\boldsymbol{v}_h^\ell\|^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2\sum_{\ell=0}^{J-1}\|\boldsymbol{\nabla}\boldsymbol{v}_h^\ell\|^2 \\
&\leq \frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|^2 + k\sum_{\ell=0}^{J-1}\langle\boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell\rangle.
\end{aligned} \tag{5.26}$$

Proof. To abbreviate notation we define

$$\boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell) := \ell_{\mathrm{ex}}^2\boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k\boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell) \in \mathcal{S}^1(\mathcal{T}_h)^3.$$

Testing (5.21) with $\boldsymbol{w}_h = \boldsymbol{v}_h^\ell$, $\boldsymbol{w}_h = \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)$, and $\boldsymbol{w}_h = \boldsymbol{\mathcal{I}}_h(\boldsymbol{m}_h^\ell \times \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell))$, respectively, leads to

$$(1+\alpha^2)\|\boldsymbol{v}_h^\ell\|_h^2 = \langle\boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h + \alpha\langle\boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h, \tag{5.27a}$$

$$\alpha\|\boldsymbol{m}_h^\ell \times \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\|_h^2 = (1+\alpha^2)\langle\boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h, \tag{5.27b}$$

$$\|\boldsymbol{m}_h^\ell \times \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\|_h^2 = (1+\alpha^2)\langle\boldsymbol{m}_h^\ell \times \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)\rangle_h, \tag{5.27c}$$

---

[1]The assumption (5.23) is usually referred to as angle condition, because in 3D it is satisfied, e.g., if all dihedral angles of all tetrahedra of $\mathcal{T}_h$ are $\leq \pi/2$.

where we used $\mathcal{I}_h(|\boldsymbol{m}_h^\ell|^2) = 1$ and $\mathcal{I}_h(\boldsymbol{m}_h^\ell \cdot \boldsymbol{v}_h^\ell) = 0$ in Ω together with the identities (5.4b)–(5.4e). Combining (5.27a)–(5.27c) gives

$$\alpha \|\boldsymbol{v}_h^\ell\|_h^2 = \langle \boldsymbol{v}_h^\ell, \boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell) \rangle_h \,.$$

Plugging in the definition of $\boldsymbol{h}_{\mathrm{eff},h}^{\mathrm{imex}}(\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell)$, we see

$$\ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{v}_h^\ell, \boldsymbol{\nabla} \boldsymbol{m}_h^\ell \rangle = -\alpha \|\boldsymbol{v}_h^\ell\|_h^2 - \ell_{\mathrm{ex}}^2 \theta k \|\boldsymbol{\nabla} \boldsymbol{v}_h^\ell\|^2 + \langle \boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell \rangle. \qquad (5.28)$$

Using the angle condition, we deduce that

$$\frac{\ell_{\mathrm{ex}}^2}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^{\ell+1}\|^2 - \frac{\ell_{\mathrm{ex}}^2}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^\ell\|^2 \overset{(5.24)}{\leq} \frac{\ell_{\mathrm{ex}}^2}{2} \|\boldsymbol{\nabla}(\boldsymbol{m}_h^\ell + k \boldsymbol{v}_h^\ell)\|^2 - \frac{\ell_{\mathrm{ex}}^2}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^\ell\|^2$$

$$= \ell_{\mathrm{ex}}^2 k \langle \boldsymbol{\nabla} \boldsymbol{m}_h^\ell, \boldsymbol{\nabla} \boldsymbol{v}_h^\ell \rangle + \frac{\ell_{\mathrm{ex}}^2}{2} k^2 \|\boldsymbol{\nabla} \boldsymbol{v}_h^\ell\|^2$$

$$\overset{(5.28)}{=} -\alpha k \|\boldsymbol{v}_h^\ell\|_h^2 - \ell_{\mathrm{ex}}^2 (\theta - 1/2) k^2 \|\boldsymbol{\nabla} \boldsymbol{v}_h^\ell\|^2 + k \langle \boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell \rangle.$$

Summing over $\ell = 0, \dots, J-1$, we obtain that

$$\frac{\ell_{\mathrm{ex}}^2}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^J\|^2 + \alpha k \sum_{\ell=0}^{J-1} \|\boldsymbol{v}_h^\ell\|_h^2 + \ell_{\mathrm{ex}}^2 (\theta - 1/2) k^2 \sum_{\ell=0}^{J-1} \|\boldsymbol{\nabla} \boldsymbol{v}_h^\ell\|^2$$

$$\leq \frac{\ell_{\mathrm{ex}}^2}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^0\|^2 + k \sum_{\ell=0}^{J-1} \langle \boldsymbol{v}_h^\ell, \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) + \boldsymbol{f}^\ell \rangle.$$

Finally, the norm equivalence (5.11) yields (5.26). $\qquad \square$

Remark 5.3.7. *(i) The stability (5.26) is the very same estimate that one obtains for the first-order tangent plane scheme from [Alo08]; see, e.g., [BFF$^+$14, Lemma 3.5]. Combining this estimate with the stability of $\boldsymbol{\pi}_h$ from (5.25), one obtains boundedness of the discrete solutions, which allows to apply the standard compactness argument for parabolic PDEs to prove convergence; see, e.g., [Alo08, Section 3] or [BFF$^+$14, Section 3.5].*
(ii) Consequently, for both Algorithm 5.3.1 and Algorithm 5.3.5, one obtains a convergence result identical to [Alo08, Theorem 2, Remark 1]. In particular, as $h, k \to 0$, for $1/2 < \theta \leq 1$ no coupling of the discretization parameters is necessary, while the CFL conditions $k = o(h)$ and $k = o(h^2)$ are proved to be sufficient for $\theta = 1/2$ and $0 \leq \theta < 1/2$, respectively.
(iii) We note that [KW18, Theorem 2.2] and its proof are slightly inaccurate and, in particular, the CFL condition $k = o(h)$ is missing for $\theta = 1/2$.

We briefly comment on a projection-free modification of `PC1+IMEX`.

Remark 5.3.8. *As pointed out in Remark 5.3.4, Algorithm 5.3.5 and the first-order tangent-plane scheme from [Alo08] coincide up to mass-lumped integration in the predictor (5.21). Hence, an obvious modification of Algorithm 5.3.5 in the spirit of the projection-free tangent-plane scheme from [AHP$^+$14, Algorithm 6] is omitting the projection in the corrector (5.22), i.e., defining $\boldsymbol{m}_h^{\ell+1} := \boldsymbol{m}_h^\ell + k \boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$. For this projection-free variant of Algorithm 5.3.5, at first glance, one could hope for the same desirable theoretical features as for*

the projection-free tangent plane scheme — namely stability and weak convergence [AHP$^+$14] without the angle condition (5.23) and even strong convergence [FT17a], both at the price of a slight deterioration from nodewise unit-length $\boldsymbol{m}_h^\ell \notin \boldsymbol{\mathcal{M}}_h$. In contrast to the projection-free tangent plane scheme, the projection-free variant of Algorithm 5.3.5 is unconditionally well-posed even for the limit case $\alpha = 0$; see Remark 5.4.6(iii)–(iv) below. Further, it satisfies a discrete energy law, which, e.g., in the exchange-only case for $\theta = 1/2$ reads

$$\frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^J\|^2 + \frac{\alpha}{1+\alpha^2}\ell_{\mathrm{ex}}^4 k \sum_{\ell=0}^{J-1}\|\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)\|_h^2 = \frac{\ell_{\mathrm{ex}}^2}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|^2 \,.$$

However, due to the loss of nodewise unit-length $\boldsymbol{m}_h^\ell \notin \boldsymbol{\mathcal{M}}_h$, equivalence of the predictor of the projection-free version of Algorithm 5.3.5 and the discrete tangent space system (5.17) in $\boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^\ell]$ does not hold anymore. Consequently, the analysis for the projection-free tangent plane scheme from [AHP$^+$14, FT17a] does not (directly) transfer, and a rigorous analysis of the projection-free version of Algorithm 5.3.5 remains open.

5.4 Second-order predictor-corrector scheme

In this section, we discuss the second-order scheme proposed in [KW18]. Our contribution is threefold: In theory, well-posedness (for the predictor) of the scheme (which was left open in [KW18]) follows already from our analysis in Section 5.3.2. When accounting for the use of inexact (iterative) linear solvers, which is inevitable in practice, however, discrete unit-length $\boldsymbol{m}_h^\ell \in \boldsymbol{\mathcal{M}}_h$ is lost and therefore a conceptually new analysis is required to guarantee well-posedness in practice. We fill this fundamental gap in the analysis of [KW18] for their second-order scheme, by proving unconditional well-posedness not only for the proposed predictor-corrector scheme, but also for its practical version incorporating inexact (iterative) linear solvers. Again, we first consider the method for the case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2\boldsymbol{\Delta}\boldsymbol{m}$. The general case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2\boldsymbol{\Delta}\boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$ is treated in Section 5.4.4, where we employ an explicit treatment of the (nonlocal) lower-order contributions to obtain a computationally superior IMEX version of the scheme, preserving experimental rates in time. We numerically confirm the applicability and the formal second-order of the proposed IMEX scheme in Section 5.5. Theoretical stability (and hence convergence) of the second-order scheme remains open (like in [KW18]), but is experimentally investigated in a numerical study in Section 5.5.3.

5.4.1 Variational formulation

The following algorithm restates [KW18, Algorithm 2] written in terms of the discrete functions $\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$. In particular, the corrector (5.30) of Algorithm 5.4.1 reformulates the N equations in $\mathbb{R}^3$ of the corrector of [KW18, Algorithm 2] as an equivalent variational formulation for $\boldsymbol{m}_h^{\ell+1}$ in $\mathcal{S}^1(\mathcal{T}_h)^3$. The predictor step coincides with step (i) of Algorithm 5.3.1, i.e., (5.29) coincides with (5.15). As in Section 5.3, the parameter $\theta \in [0,1]$ modulates the 'degree of implicitness' (in the predictor) of the scheme.

Algorithm 5.4.1 (PC2, variational form). ***Input:*** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$.
Loop: *For all time-steps $\ell = 0, \dots, L-1$, iterate:*

(i) *Compute* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$(1 + \alpha^2)\langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h = -\ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell), \boldsymbol{w}_h \rangle_h$$
$$- \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell)), \boldsymbol{w}_h \rangle_h \,. \tag{5.29}$$

(ii) *Compute* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ *such that, for all* $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, *it holds that*

$$(1 + \alpha^2)\langle d_t \boldsymbol{m}_h^{\ell+1}, \boldsymbol{w}_h \rangle_h = -\ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^{\ell+1/2} \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell), \boldsymbol{w}_h \rangle_h \tag{5.30}$$
$$- \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^{\ell+1/2} \times [(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)], \boldsymbol{w}_h \rangle_h \,.$$

Output: *Sequence of discrete functions* $\left\{ (\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1}) \right\}_{\ell=0}^{L-1}$.

The corrector step of Algorithm 5.3.1, which combines a linear first-order time-stepping with the nodal projection, is replaced by the linear system (5.30). The 2D numerical results of [KW18, Figure 3] indicate that the method is of second-order in time. In Section 5.5.2, we confirm this observation for a numerical example in 3D.

5.4.2 Unconditional well-posedness, exact solver

In Lemma 5.4.2, we first collect two basic properties of Algorithm 5.4.1, which, for $\alpha > 0$, turn out to be sufficient to prove unconditional well-posedness of the algorithm in Theorem 5.4.3.

Lemma 5.4.2. *Let* $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$. *Suppose that the solutions* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ *and* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$ *to* (5.29) *and* (5.30) *exist, respectively. Then,* $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, *and* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.

Proof. The claim $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ follows as in the proof of Lemma 5.3.2. We show that $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ implies $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ due to the corrector system (5.30): For arbitrary $\boldsymbol{z} \in \mathcal{N}_h$, we choose $\boldsymbol{w}_h := \boldsymbol{m}_h^{\ell+1/2}(\boldsymbol{z})\phi_{\boldsymbol{z}} \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (5.30) to see

$$\frac{(1 + \alpha^2)\beta_{\boldsymbol{z}}}{2k}\left(|\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z})|^2 - |\boldsymbol{m}_h^\ell(\boldsymbol{z})|^2 \right) \overset{(5.10)}{=} (1 + \alpha^2)\langle d_t \boldsymbol{m}_h^{\ell+1}, \boldsymbol{m}_h^{\ell+1/2}(\boldsymbol{z})\phi_{\boldsymbol{z}} \rangle_h \overset{(5.30),(5.4b)}{=} 0\,.$$

This shows that $|\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z})| = |\boldsymbol{m}_h^\ell(\boldsymbol{z})|$ for all $\boldsymbol{z} \in \mathcal{N}_h$. Hence, $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ implies that $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$. The assumption $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ concludes the proof. $\square$

We show unconditional well-posedness of the corrector (5.30), while with Lemma 5.4.2 unconditional well-posedness of the predictor is inferred from our analysis in Section 5.3.2.

Theorem 5.4.3. *Let* $\alpha > 0$. *Then, Algorithm 5.4.1 is unconditionally well-posed for any input* $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, *i.e., for all* $\ell = 0, \dots, L-1$, *the predictor* (5.29) *admits a unique solution* $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$, *and the corrector* (5.30) *admits a unique solution* $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.

Proof. By Lemma 5.4.2 it holds that $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ and $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$ for all $\ell = 0, \dots, L-1$. Hence, as for the predictor of Algorithm 5.3.1, the predictor system (5.29) is equivalent to a coercive system in the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$ with unique solution $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$;

see (the proof of) Theorem 5.3.3. It remains to show well-posedness of the corrector (5.30): We rewrite the problem in terms of the unknown $\boldsymbol{\eta}_h^\ell := \boldsymbol{m}_h^{\ell+1/2}$, which, by construction, satisfies that $\boldsymbol{m}_h^{\ell+1} = 2\boldsymbol{\eta}_h^\ell - \boldsymbol{m}_h^\ell$ and $d_t\boldsymbol{m}_h^{\ell+1} = 2(\boldsymbol{\eta}_h^\ell - \boldsymbol{m}_h^\ell)/k$. The corrector system (5.30) then reads: Find $\boldsymbol{\eta}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that

$$a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell](\boldsymbol{\eta}_h^\ell, \boldsymbol{w}_h) = (1 + \alpha^2)\langle \boldsymbol{m}_h^\ell, \boldsymbol{w}_h\rangle_h,$$

where the bilinear form $a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell] \colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$ is defined by

$$a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell](\boldsymbol{\eta}_h^\ell, \boldsymbol{w}_h) := (1 + \alpha^2)\langle \boldsymbol{\eta}_h^\ell, \boldsymbol{w}_h\rangle_h + \frac{\ell_{\mathrm{ex}}^2 k}{2} \langle \boldsymbol{\eta}_h^\ell \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell), \boldsymbol{w}_h\rangle_h$$
$$+ \frac{\alpha\ell_{\mathrm{ex}}^2 k}{2} \langle \boldsymbol{\eta}_h^\ell \times [(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) \times \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)], \boldsymbol{w}_h\rangle_h.$$

As the bilinear form satisfies the ellipticity property

$$a_{\mathrm{cor}}[\boldsymbol{m}_h^\ell, \boldsymbol{v}_h^\ell](\boldsymbol{w}_h, \boldsymbol{w}_h) = (1 + \alpha^2)\|\boldsymbol{w}_h\|_h^2 \quad \text{for all } \boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3,$$

the problem is well-posed by the Lax–Milgram theorem. Hence, (5.30) provides a unique solution $\boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$. Lemma 5.4.2 guarantees $\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$ concluding the proof. $\qquad\square$

Remark 5.4.4. *(i) Algorithm 5.4.1 is a predictor-corrector scheme: Both systems, for the predictor (5.29) and for the corrector (5.30), respectively, are linear systems representing discrete mass-lumped variational versions of the LL form (5.5a) of LLG; see also Remark 5.3.4(i). First, treating the effective field implicitly in time, an approximate time derivative $\boldsymbol{v}_h^\ell \in \mathcal{K}_h[\boldsymbol{m}_h^\ell]$, the predictor, is computed. In the second step (the effective field of) the predicted midpoint $\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ is used to compute a corrected update $d_t\boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$, guaranteeing conservation of discrete unit-length $\boldsymbol{m}_h^{\ell+1} := \boldsymbol{m}_h^\ell + kd_t\boldsymbol{m}_h^{\ell+1} \in \mathcal{M}_h$.*
(ii) In the proof of Theorem 5.4.3, note that the assumption $\alpha > 0$ is only exploited to apply Theorem 5.3.3. Hence, analogously to Theorem 5.3.3 (Remark 5.3.4(iv)), also Theorem 5.4.3 can be extended to the limit case $\alpha = 0$; see Theorem 5.4.5 below.

5.4.3 Unconditional well-posedness, inexact solver

Considering the effect of numerical approximations, we extend the theoretical well-posedness result from the previous section to the practical case.

Well-posedness of the predictor step (i) of Algorithm 5.4.1 is guaranteed by Theorem 5.3.3: There, under the crucial condition $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$, computing $\boldsymbol{v}_h^\ell$ in the predictor step is shown to be equivalent to solving the system (5.17) in the discrete tangent space $\mathcal{K}_h[\boldsymbol{m}_h^\ell]$, which is always well-posed for $\alpha > 0$. While $\boldsymbol{m}_h^\ell \in \mathcal{M}_h$ is explicitly enforced in step (ii) of Algorithm 5.3.1, in Algorithm 5.4.1 it follows only implicitly from the inherent length preservation guaranteed by the variational formulation (5.30) solved in step (ii) together with $\boldsymbol{m}_h^{\ell-1} \in \mathcal{M}_h$ in the previous time-step; see the proof of Lemma 5.4.2. In practice however, linear systems are solved by inexact (iterative) numerical solvers, i.e., the coefficient vector of the unknown $\boldsymbol{m}_h^{\ell+1}$ solves the linear system of equations corresponding to (5.30) only up to some accuracy $\varepsilon > 0$, commonly in the $\ell^2(\mathbb{R}^{3N})$-norm. Consequently, for any $\boldsymbol{z} \in \mathcal{N}_h$

there only holds $|\boldsymbol{m}_h^{\ell+1}(\boldsymbol{z})| \approx |\boldsymbol{m}_h^\ell(\boldsymbol{z})|$ with a small error depending on the discretization parameters ε and h. Moreover, the deviation from nodewise unit-length accumulates over the time-steps $\ell = 0, \dots, L-1$. Consequently — if recoverable at all — one expects to require CFL-type couplings of the discretization parameters k, h, ε to rigorously argue (approximate) equivalence of the linear system in step (i) of Algorithm 5.4.1 and the well-posed system (5.17) in the proof of Theorem 5.3.3.

To avoid these analytical difficulties, we take a different analytical approach: The new analysis uses a space decomposition technique reformulating (5.29) as an equivalent saddle-point problem, which subsequently is proved to be unconditionally well-posed and hence always provides a unique solution. In particular, this does not require $\boldsymbol{m}_h^\ell \in \boldsymbol{\mathcal{M}}_h$, but allows for arbitrary $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3 \supsetneq \boldsymbol{\mathcal{M}}_h$. Additionally, the analysis applies to all $\alpha \geq 0$, extending well-posedness of Algorithm 5.4.1 to the Schrödinger map equation ($\alpha = 0$).

Theorem 5.4.5. *Let $\alpha \geq 0$. Then, Algorithm 5.4.1 is unconditionally well-posed for any input $\boldsymbol{m}_h^0 \in \mathcal{S}^1(\mathcal{T}_h)^3$, i.e., for all $\ell = 0, \dots, L-1$ and any $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$, the predictor (5.29) admits a unique solution $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$, and the corrector (5.30) admits a unique solution $\boldsymbol{m}_h^{\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$.*

Proof. For arbitrary $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ well-posedness of the corrector (5.30) is guaranteed by the proof of Theorem 5.4.3, as it does not require $\boldsymbol{m}_h^\ell \in \boldsymbol{\mathcal{M}}_h$. Using a space decomposition technique, we show unconditional well-posedness of the predictor system (5.29) for any $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ — in particular for $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ not necessarily belonging to $\boldsymbol{\mathcal{M}}_h$ — in five steps:

- **Step 0:** Some notation.

Throughout, for an operator $A \colon X \to Y$ between two Hilbert spaces, we write $\mathcal{R}(A) \subseteq Y$ for its range, and $\mathcal{N}(A) \subseteq X$ for its kernel. We consider the (negative) discrete Laplace operator (5.12) restricted to $\mathcal{S}^1(\mathcal{T}_h)^3 \subset \boldsymbol{H}^1(\Omega)$, which will be denoted by the same symbol $-\boldsymbol{\Delta}_h \colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathcal{S}^1(\mathcal{T}_h)^3$. Further, we identify a 3-vector with the corresponding constant vector-valued grid function, i.e., $\mathbb{R}^3 \subset \left(\mathcal{S}^1(\mathcal{T}_h)^3, \langle \cdot, \cdot \rangle_h\right)$. For $S \subset \mathcal{S}^1(\mathcal{T}_h)^3$ a subspace we denote by $\boldsymbol{I}_S$ the identity on S.

- **Step 1:** Orthogonal decomposition $\mathcal{S}^1(\mathcal{T}_h)^3 = \mathcal{R}(\boldsymbol{P}_*) \oplus \mathcal{N}(\boldsymbol{P}_*)$.

Define the operator $\boldsymbol{P}_* \colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathcal{S}^1(\mathcal{T}_h)^3$ for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$ via

$$(\boldsymbol{P}_* \boldsymbol{w}_h)_j = (\boldsymbol{w}_h)_j - \mathrm{meas}(\Omega)^{-1} \langle \boldsymbol{w}_h, \boldsymbol{e}_j \rangle_h \in \mathcal{S}^1(\mathcal{T}_h) \qquad \text{for all} \quad j = 1, 2, 3 \,.$$

Clearly, $\boldsymbol{P}_*$ is the $\langle \cdot, \cdot \rangle_h$-orthogonal projector onto

$$\mathcal{R}(\boldsymbol{P}_*) = \mathcal{S}_*^1(\mathcal{T}_h)^3 := \{\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3 \colon \langle \boldsymbol{w}_h, \boldsymbol{e}_j \rangle_h = 0 \text{ for all } j = 1, 2, 3\} \,,$$

the subset of $\mathcal{S}^1(\mathcal{T}_h)^3$ consisting of the vector-valued grid functions which have zero mean in each component. Due to self-adjointness, $\boldsymbol{P}_*$ provides the orthogonal decomposition

$$\mathcal{S}^1(\mathcal{T}_h)^3 = \mathcal{R}(\boldsymbol{P}_*) \oplus \mathcal{N}(\boldsymbol{P}_*) = \mathcal{S}_*^1(\mathcal{T}_h)^3 \oplus \mathbb{R}^3 \,.$$

With respect to this decomposition, we rewrite the unknown $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ as the orthogonal sum

$$\boldsymbol{v}_h^\ell = \boldsymbol{P}_* \boldsymbol{v}_h^\ell \oplus (\boldsymbol{I}_{\mathcal{S}^1(\mathcal{T}_h)^3} - \boldsymbol{P}_*) \boldsymbol{v}_h^\ell =: \boldsymbol{v}_* \oplus \overline{\boldsymbol{v}} \,, \tag{5.31}$$

with unique $\boldsymbol{v}_* \in \mathcal{R}(\boldsymbol{P}_*) = \mathcal{S}_*^1(\mathcal{T}_h)^3$ and $\overline{\boldsymbol{v}} \in \mathcal{N}(\boldsymbol{P}_*) = \mathbb{R}^3$. Note that $\overline{\boldsymbol{v}} \in \mathbb{R}^3$ is the vector-valued mean of $\boldsymbol{v}_h^\ell$, i.e., $\langle \overline{\boldsymbol{v}}, \boldsymbol{e}_j \rangle_h = \langle \boldsymbol{v}_h^\ell, \boldsymbol{e}_j \rangle_h$ for all components $j = 1, 2, 3$.

- **Step 2:** Reduced operator $-\widetilde{\boldsymbol{\Delta}}_h \colon \mathcal{S}_*^1(\mathcal{T}_h)^3 \to \mathcal{S}_*^1(\mathcal{T}_h)^3$.

The discrete Laplacian $-\boldsymbol{\Delta}_h \colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathcal{S}^1(\mathcal{T}_h)^3$ is linear, self-adjoint and by definition (5.12) has the kernel $\mathcal{N}(-\boldsymbol{\Delta}_h) = \mathbb{R}^3 \subset \mathcal{S}^1(\mathcal{T}_h)^3$. Hence, there holds the orthogonal decomposition

$$\mathcal{S}^1(\mathcal{T}_h)^3 = \mathcal{R}(-\boldsymbol{\Delta}_h) \oplus \mathcal{N}(-\boldsymbol{\Delta}_h) = \mathcal{N}(-\boldsymbol{\Delta}_h)^\perp \oplus \mathcal{N}(-\boldsymbol{\Delta}_h) = \mathcal{S}_*^1(\mathcal{T}_h)^3 \oplus \mathbb{R}^3 \,. \tag{5.32}$$

Consequently, the reduced operator $-\boldsymbol{\Delta}_h|_{\mathcal{S}^1(\mathcal{T}_h)^3} =: -\widetilde{\boldsymbol{\Delta}}_h \colon \mathcal{S}_*^1(\mathcal{T}_h)^3 \to \mathcal{S}_*^1(\mathcal{T}_h)^3$ is linear, self-adjoint, and bijective. Moreover, it provides a well-defined inverse denoted by $(-\widetilde{\boldsymbol{\Delta}}_h)^{-1} \colon \mathcal{S}_*^1(\mathcal{T}_h)^3 \to \mathcal{S}_*^1(\mathcal{T}_h)^3$ with the same attributes. We point out the identities

$$(-\widetilde{\boldsymbol{\Delta}}_h)^{-1} \circ (-\boldsymbol{\Delta}_h) = \boldsymbol{P}_* \quad \text{and} \quad -\boldsymbol{\Delta}_h \circ (-\widetilde{\boldsymbol{\Delta}}_h)^{-1} = \boldsymbol{P}_*|_{\mathcal{S}_*^1(\mathcal{T}_h)^3} = \boldsymbol{I}_{\mathcal{S}_*^1(\mathcal{T}_h)^3} \,, \tag{5.33}$$

which follow from the orthogonal decomposition (5.32).

- **Step 3:** Equivalent saddle point formulation.

With the unknowns $\boldsymbol{q} := -\boldsymbol{\Delta}_h \boldsymbol{v}_* \in \mathcal{S}_*^1(\mathcal{T}_h)^3$ and $\boldsymbol{\lambda} := \overline{\boldsymbol{v}} \in \mathbb{R}^3$ from (5.31), we induce the representation $\boldsymbol{v}_h^\ell = (-\widetilde{\boldsymbol{\Delta}}_h)^{-1} \boldsymbol{q} \oplus \boldsymbol{\lambda}$. Plugging this identity into (5.29), we rewrite the predictor as equivalent saddle point problem: Find $(\boldsymbol{q}, \boldsymbol{\lambda}) \in \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3$, such that for all $(\boldsymbol{w}, \boldsymbol{\mu}) \in \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3$ it holds that

$$a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{q}, \boldsymbol{w}) + b_{\mathrm{sp}}(\boldsymbol{w}, \boldsymbol{\lambda}) = F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}) \,, \tag{5.34a}$$

$$b_{\mathrm{sp}}(\boldsymbol{q}, \boldsymbol{\mu}) = 0 \,, \tag{5.34b}$$

with the (bi-)linear forms $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell] \colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$, $b \colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3 \to \mathbb{R}$, and $F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell] \colon \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$ given by

$$\begin{aligned}
a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{q}, \boldsymbol{w}) &:= (1 + \alpha^2) \langle (-\widetilde{\boldsymbol{\Delta}}_h)^{-1} \boldsymbol{P}_* \boldsymbol{q}, \boldsymbol{w} \rangle_h \\
&\quad - \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{m}_h^\ell \times \boldsymbol{q}, \boldsymbol{w} \rangle_h - \alpha \ell_{\mathrm{ex}}^2 \theta k \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{q}), \boldsymbol{w} \rangle_h \,, \\
b_{\mathrm{sp}}(\boldsymbol{w}, \boldsymbol{\lambda}) &:= (1 + \alpha^2) \langle \boldsymbol{\lambda}, \boldsymbol{w} \rangle_h \,, \\
F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{w}) &:= -\ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{m}_h^\ell, \boldsymbol{w} \rangle_h - \alpha \ell_{\mathrm{ex}}^2 \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times \boldsymbol{\Delta}_h \boldsymbol{m}_h^\ell), \boldsymbol{w} \rangle_h \,.
\end{aligned}$$

The equivalence of (5.34a)–(5.34b) to (5.29) follows from $\boldsymbol{\lambda} \in \mathcal{N}(-\boldsymbol{\Delta}_h)$ and (5.33). We use the operator $(-\widetilde{\boldsymbol{\Delta}}_h)^{-1} \circ \boldsymbol{P}_*$ rather than $(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}$, so that the bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ is well-defined on $\mathcal{S}^1(\mathcal{T}_h)^3 \supsetneq \mathcal{S}_*^1(\mathcal{T}_h)^3$. The second equation (5.34b) ensures $\boldsymbol{q} \in \mathcal{S}_*^1(\mathcal{T}_h)^3$, which is not enforced explicitly.

- **Step 4:** The bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ is coercive on the kernel of b_{sp}.

We aim to apply the Brezzi theory for saddle point problems; see, e.g., [BBF13, Section 4.2]. Hence, we require coercivity of the bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell] \colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3 \to \mathbb{R}$ on

$$\bigcap_{\boldsymbol{\lambda}\in\mathbb{R}^3} \mathcal{N}\big(b_{\mathrm{sp}}(\cdot,\boldsymbol{\lambda})\big) = \bigcap_{\boldsymbol{\lambda}\in\mathbb{R}^3} \{\boldsymbol{w}\in\mathcal{S}^1(\mathcal{T}_h)^3 \colon \langle\boldsymbol{\lambda},\boldsymbol{w}\rangle_h = 0\}$$
$$= \bigcap_{j=1,2,3} \{\boldsymbol{w}\in\mathcal{S}^1(\mathcal{T}_h)^3 \colon \langle\boldsymbol{e}_j,\boldsymbol{w}\rangle_h = 0\} = \mathcal{S}_*^1(\mathcal{T}_h)^3\,.$$

For any $\boldsymbol{q}\in\mathcal{S}_*^1(\mathcal{T}_h)^3$, we compute

$$a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell](\boldsymbol{q},\boldsymbol{q}) \overset{(5.4\mathrm{b})}{=} (1+\alpha^2)\langle(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{P}_*\boldsymbol{q},\boldsymbol{q}\rangle_h - \alpha\ell_{\mathrm{ex}}^2\theta k\langle\boldsymbol{m}_h^\ell\times(\boldsymbol{m}_h^\ell\times\boldsymbol{q}),\boldsymbol{q}\rangle_h\,,$$
$$\overset{(5.33),(5.4\mathrm{d})}{=} (1+\alpha^2)\langle(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q},-\boldsymbol{\Delta}_h(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\rangle_h + \alpha\ell_{\mathrm{ex}}^2\theta k\|\boldsymbol{m}_h^\ell\times\boldsymbol{q}\|_h^2$$
$$\overset{(5.12)}{=} (1+\alpha^2)\|\boldsymbol{\nabla}(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\|_{\boldsymbol{L}^2(\Omega)}^2 + \alpha\ell_{\mathrm{ex}}^2\theta k\|\boldsymbol{m}_h^\ell\times\boldsymbol{q}\|_h^2$$
$$\gtrsim h^2\|\boldsymbol{q}\|_h^2 + \alpha\ell_{\mathrm{ex}}^2\theta k\|\boldsymbol{m}_h^\ell\times\boldsymbol{q}\|_h^2 \geq h^2\|\boldsymbol{q}\|_h^2\,,$$

where the second to last estimate is an inverse estimate on $\mathcal{S}_*^1(\mathcal{T}_h)^3$ derived from the classical inverse estimate on $\mathcal{S}^1(\mathcal{T}_h)^3$ via

$$\|\boldsymbol{q}\|_h^2 = \langle\boldsymbol{q},\boldsymbol{q}\rangle_h \overset{(5.33)}{=} \langle\boldsymbol{q},-\boldsymbol{\Delta}_h(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\rangle_h \overset{(5.12)}{=} \langle\boldsymbol{\nabla}\boldsymbol{q},\boldsymbol{\nabla}(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\rangle_{\boldsymbol{L}^2(\Omega)}$$
$$\leq \|\boldsymbol{\nabla}\boldsymbol{q}\|_{\boldsymbol{L}^2(\Omega)}\|\boldsymbol{\nabla}(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\|_{\boldsymbol{L}^2(\Omega)} \lesssim h^{-1}\|\boldsymbol{q}\|_h\|\boldsymbol{\nabla}(-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\|_{\boldsymbol{L}^2(\Omega)}\,.$$

Hence, $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ is coercive on $\bigcap_{\boldsymbol{\lambda}\in\mathbb{R}^3}\mathcal{N}(b_{\mathrm{sp}}(\cdot,\boldsymbol{\lambda})) = \mathcal{S}_*^1(\mathcal{T}_h)^3$ with ellipticity constant proportional to $h^2 > 0$.

- **Step 5:** Unique solvability and reconstruction of $\boldsymbol{v}_h^\ell$.

Clearly, $b_{\mathrm{sp}}\colon \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathbb{R}^3 \to \mathbb{R}$ satisfies the inf-sup condition with constant $(1+\alpha^2) > 0$. Now unique solvability of the saddle point formulation (5.34a)–(5.34b) follows from the Brezzi theorem [BBF13, Theorem 4.2.1]. Ultimately, with $(\boldsymbol{q},\boldsymbol{\lambda})\in\mathcal{S}_*^1(\mathcal{T}_h)^3\times\mathbb{R}^3$ denoting the unique solution of (5.34a)–(5.34b), the original unknown solution to (5.29) is reconstructed via $\boldsymbol{v}_h^\ell = (-\widetilde{\boldsymbol{\Delta}}_h)^{-1}\boldsymbol{q}\oplus\boldsymbol{\lambda}\in\mathcal{S}_*^1(\mathcal{T}_h)^3\oplus\mathbb{R}^3 = \mathcal{S}^1(\mathcal{T}_h)^3$ and is therefore also unique. □

Remark 5.4.6. *(i) In the third step of the proof of Theorem 5.4.5, we introduced the unknown $\boldsymbol{q} := -\boldsymbol{\Delta}_h\boldsymbol{v}_*\in\mathcal{S}_*^1(\mathcal{T}_h)^3$. This idea is inspired by [XGCW⁺20, Section 2.3], where the authors subsequently use the Browder–Minty lemma for monotone operators to prove well-posedness of their proposed finite difference LLG integrator based on the second-order backward differentiation formula.*
(ii) In Step 4 of the proof of Theorem 5.4.5, as the new unknown $\boldsymbol{q} = -\boldsymbol{\Delta}_h\boldsymbol{v}_h^\ell$ comprises second-order derivatives of the original unknown, it is not surprising that the ellipticity constant for the bilinear form $a_{\mathrm{sp}}[\boldsymbol{m}_h^\ell]$ scales proportionally to $h^2 > 0$.
(iii) Since $\boldsymbol{\mathcal{M}}_h\subset\mathcal{S}^1(\mathcal{T}_h)^3$ and the predictors of Algorithm 5.3.1 and Algorithm 5.4.1 coincide, the proof of Theorem 5.4.5 is not only an alternative proof to Theorem 5.4.3, but also to Theorem 5.3.3, which additionally extends both theorems to the critical value $\alpha = 0$.
(iv) Consequently, Algorithm 5.3.1 is not only a mass-lumped version of the tangent plane

scheme [Alo08], but additionally it is well-posed for the Schrödinger map equation ($\alpha = 0$).
(v) Even though the predictor of Algorithm 5.3.1 written in the form (5.17) *coincides with the predictor of the tangent plane scheme up to the used integration rule, well-posedness of the tangent plane scheme for the limit case $\alpha = 0$ remains open. Indeed, the proof of Theorem 5.4.5 relies heavily on mass-lumped integration, and we did not succeed to transfer the proof to exact integration used in the original tangent plane scheme.*

5.4.4 Including the lower-order contributions

We consider the case when the effective field comprises linear lower-order energy contributions $\boldsymbol{\pi}(\boldsymbol{m})$ such as, in particular, the nonlocal stray field $\boldsymbol{h}_{\mathrm{s}}$, i.e., $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \boldsymbol{\Delta} \boldsymbol{m} + \boldsymbol{\pi}(\boldsymbol{m}) + \boldsymbol{f}$. Then the predictor step of the original second-order integrator proposed in [KW18, Algorithm 2] is identical to (5.20), i.e., lower-order terms are treated implicitly in time. Due to the nonlocality of the stray field this is unattractive in practice as described in Section 5.3.3. Hence, analogously to Section 5.3.3, we aim to treat the lower-order terms $\boldsymbol{\pi}(\boldsymbol{m})$ explicitly in time. However, to avoid spoiling the scheme's potential second-order accuracy in time, which was observed experimentally in [KW18], the modification is slightly more involved:

In Section 5.3.3 an error of order $\mathcal{O}(k)$ is introduced to the system (5.20) by approximating $\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) \approx \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell)$. Since Algorithm 5.3.1 is a first-order scheme, this modification did not deteriorate the order of convergence of the algorithm.

To preserve the potential second-order of Algorithm 5.4.1, we use a higher-order approximation to $\boldsymbol{\pi}(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell)$: Recall that $\boldsymbol{\pi}$ is a linear operator and that $\boldsymbol{v}_h^\ell$ is an approximation of $\partial_t \boldsymbol{m}(t_\ell)$. Motivated by the Taylor expansion $\boldsymbol{m}(t_\ell) = \boldsymbol{m}(t_{\ell-1}) + k \partial_t \boldsymbol{m}(t_\ell) + \mathcal{O}(k^2)$, and hence $\boldsymbol{m}(t_\ell) + \theta k \partial_t \boldsymbol{m}(t_\ell) = (1 + \theta) \boldsymbol{m}(t_\ell) - \theta \boldsymbol{m}(t_{\ell-1}) + \mathcal{O}(k^2)$, we introduce a second-order error $\mathcal{O}(k^2)$ to the system (5.20) via the approximation

$$\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) \approx (1 + \theta) \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta \boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1}) \,.$$

Only the leading-order exchange contribution is treated implicitly in time, while the lower-order contributions are treated explicitly in time. Due to the higher-order approximation of $\boldsymbol{\pi}_h(\boldsymbol{v}_h^\ell)$, this does not spoil the observed second-order of the scheme and it is computationally much more attractive. To sum up, we consider the following algorithm.

Algorithm 5.4.7 (`PC2+IMEX`)**.** ***Input:*** $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$.
Preprocessing: *Compute $\boldsymbol{m}_h^1 \in \boldsymbol{\mathcal{M}}_h$, e.g., by Algorithm 5.4.1.*
Loop: *For all time-steps $\ell = 1, \dots, L - 1$, iterate:*

(i) *Compute $\mathbb{P}_h((1 + \theta) \boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta \boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1}) + \boldsymbol{f}^{\ell+\theta}) \in \mathcal{S}^1(\mathcal{T}_h)^3$.*

(ii) *Compute $\boldsymbol{v}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that*

$$\begin{aligned}
(1 &+ \alpha^2) \langle \boldsymbol{v}_h^\ell, \boldsymbol{w}_h \rangle_h \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (5.35) \\
&= -\langle \boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h((1+\theta)\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta \boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1}) + \boldsymbol{f}^{\ell+\theta})], \boldsymbol{w}_h \rangle_h \\
&\quad - \alpha \langle \boldsymbol{m}_h^\ell \times (\boldsymbol{m}_h^\ell \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + \theta k \boldsymbol{v}_h^\ell) + \mathbb{P}_h((1+\theta)\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell) - \theta \boldsymbol{\pi}_h(\boldsymbol{m}_h^{\ell-1}) + \boldsymbol{f}^{\ell+\theta})]), \boldsymbol{w}_h \rangle_h.
\end{aligned}$$

(iii) Compute $\boldsymbol{m}_h^{\ell+1} \in \boldsymbol{\mathcal{M}}_h$ such that, for all $\boldsymbol{w}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that

$$
\begin{aligned}
(1 + \alpha^2)\langle d_t \boldsymbol{m}_h^{\ell+1}, \boldsymbol{w}_h\rangle_h \\
= -\langle \boldsymbol{m}_h^{\ell+1/2} \times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+1/2})], \boldsymbol{w}_h\rangle_h \\
- \alpha\langle \boldsymbol{m}_h^{\ell+1/2} \times \big((\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) \\
\times [\ell_{\mathrm{ex}}^2 \boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \mathbb{P}_h(\boldsymbol{\pi}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell) + \boldsymbol{f}^{\ell+1/2})]\big), \boldsymbol{w}_h\rangle_h.
\end{aligned}
$$

Output: *Sequence of discrete functions $\left\{(\boldsymbol{v}_h^\ell, \boldsymbol{m}_h^{\ell+1})\right\}_{\ell=0}^{L-1}$.*

Remark 5.4.8. *(i) In the preprocessing step of Algorithm 5.4.7 also other integrators may be used to compute $\boldsymbol{m}_h^1 \in \boldsymbol{\mathcal{M}}_h$. As long as the approximation $\boldsymbol{m}_h^1$ is second-order accurate, the potential second-order accuracy of Algorithm 5.4.1 is preserved by Algorithm 5.4.7. (Note that first-order accurate integrators usually only introduce a quadratic error per time-step.) (ii) Algorithm 5.4.7 is also well-posed in practice, when effects of inexact (iterative) solvers are accounted for, i.e., (5.35) is unconditionally well-posed for arbitrary $\boldsymbol{m}_h^\ell \in \mathcal{S}^1(\mathcal{T}_h)^3 \supsetneq \boldsymbol{\mathcal{M}}_h$. As lower-order terms are treated explicitly in time, proving well-posedness follows the lines of the proof of Theorem 5.4.5 with adjusted linear form $F_{\mathrm{sp}}[\boldsymbol{m}_h^\ell] \rightsquigarrow F_{\mathrm{imex}}[\boldsymbol{m}_h^\ell, \boldsymbol{m}_h^{\ell-1}]$.*

5.5 Numerical experiments

This section provides some numerical experiments for Algorithm 5.3.1 and Algorithm 5.4.1 from [KW18], as well as their respective IMEX versions proposed in this work, namely Algorithm 5.3.5 and Algorithm 5.4.7, respectively. In Section 5.5.1 we verify the correctness of the proposed integrators (`PC1+IMEX` and `PC2+IMEX`) on the benchmark problem μMAG #4 from [MUM]. In Section 5.5.2 the experimental rates of Algorithm 5.3.1 (`PC1`) and Algorithm 5.4.1 (`PC2`) reported in [KW18] are confirmed. Moreover, the experiment shows that lower-order terms can appropriately be treated explicitly in time by Algorithm 5.3.5 (`PC1+IMEX`) and Algorithm 5.4.7 (`PC2+IMEX`), respectively, without spoiling the rate of convergence.

All computations have been performed with our micromagnetic software module Commics [PRS$^+$20], based on the open-source finite element library Netgen/NGSolve [ngs]. In Commics, the stray field $\boldsymbol{h}_{\mathrm{s}}$ is computed via the hybrid FEM-BEM approach from [FK90]. We note that meshes generated by Netgen in general do not satisfy the angle condition (5.23). All experiments were repeated on structured meshes satisfying the angle condition leading to the same results (not displayed).

5.5.1 μMAG standard problem #4

We verify the practical applicability of the proposed integrators `PC1+IMEX` and `PC2+IMEX` (we choose $\theta = 1/2$) by computing a physically relevant example. To this end, we consider μMAG standard problem #4 [MUM], which simulates the switching of the magnetization in a thin permalloy layer.

The objective is the simulation of the magnetization dynamics in a thin permalloy film of dimensions $500\,\mathrm{nm} \times 125\,\mathrm{nm} \times 3\,\mathrm{nm}$ under the influence of a constant applied external

field. The involved physical constants and material parameters are the gyromagnetic ratio $\gamma_0 = 2.211 \cdot 10^5 \, \mathrm{m/C}$, the permeability of vacuum $\mu_0 = 4\pi \cdot 10^{-7} \, \mathrm{N/A^2}$, the saturation magnetization $M_\mathrm{s} = 8.0 \cdot 10^5 \, \mathrm{A/m}$, the exchange stiffness constant $A = 1.3 \cdot 10^{-11} \, \mathrm{J/m}$, and the Gilbert damping constant $\alpha = 0.02$. Starting from a so-called equilibrium S-state [MUM], the experiment consists in applying the constant applied field $\mu_0 \boldsymbol{H}_\mathrm{ext} = (-24.6, 4.3, 0) \, \mathrm{mT}$ for $3 \, \mathrm{ns}$.

For the rescaled form (5.5) of LLG, the above physical quantities lead to the parameters $\ell_\mathrm{ex} = \sqrt{2A/(\mu_0 M_\mathrm{s}^2)}$, $T = 3 \cdot 10^{-9} \gamma_0 M_\mathrm{s}$, and $\boldsymbol{f} = \boldsymbol{H}_\mathrm{ext}/M_\mathrm{s}$, while $\boldsymbol{\pi}(\boldsymbol{m})$ includes only the stray field $\boldsymbol{h}_\mathrm{s}$. For the space discretization, we consider a tetrahedral partition of the thin film generated by Netgen [ngs] into cells of prescribed mesh size $3 \, \mathrm{nm}$. This corresponds to $48\,796$ elements and $16\,683$ vertices. For the time discretization, we consider a constant physical time-step size of $\Delta t = 0.1 \, \mathrm{ps}$, which is connected to the rescaled time-step size k via the relation $k = \gamma_0 M_\mathrm{s} \Delta t$.

For comparison, the desired output of this benchmark problem is the evolution of the x-, y- and z-component of the spatially averaged magnetization. Figure 5.1 shows that our results match those computed by the finite difference code OOMMF [OOM] available on the μMAG homepage [MUM].
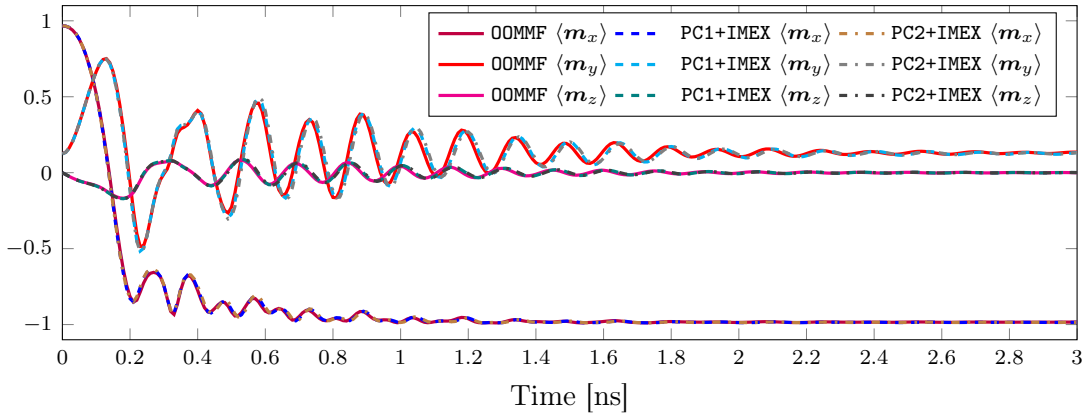
Figure 5.1: μMAG standard problem #4 from Section 5.5.1: Time evolution of the spatially averaged magnetization components computed with Algorithm 5.3.5 (`PC1+IMEX`) and Algorithm 5.4.7 (`PC2+IMEX`) compared to the results of `OOMMF`.

5.5.2 Empirical convergence rates for LLG

We aim to illustrate the accuracy and the computational effort of the following four algorithms:

- `PC1`: fully implicit first-order scheme proposed in [KW18] and recalled in Algorithm 5.3.1;

- `PC1+IMEX`: `PC1` with explicit treatment of the lower-order terms as proposed in this work and formulated in Algorithm 5.3.5;

- `PC2`: fully implicit second-order scheme proposed in [KW18] and recalled in Algorithm 5.4.1;

- `PC2+IMEX`: `PC2` with explicit treatment of the lower-order terms as proposed in this work and formulated in Algorithm 5.4.7;

For all integrators we choose $\theta = 1/2$. To obtain experimental convergence rates in time, we use the model problem proposed in [PRS18b]: We consider the initial boundary value problem (5.5) with $\Omega = (0,1)^3$, $\boldsymbol{m}^0 \equiv (1,0,0)$, $\alpha = 1$, and $T = 5$. For the effective field (5.5d), we choose $\ell_{\mathrm{ex}} = 1$, a constant applied field $\boldsymbol{f} \equiv (-2, -0.5, 0)$, as well as an operator $\boldsymbol{\pi}$ which consists only of the stray field, i.e., $\boldsymbol{\pi}(\boldsymbol{m}) = \boldsymbol{h}_{\mathrm{s}}(\boldsymbol{m})$.

For the predictor step in `PC1` and `PC2`, respectively, we solve (5.20). Since $\boldsymbol{\pi}_h$ effectively depends on $\boldsymbol{v}_h^\ell$ in (5.20), the linear system in the predictor step of Algorithm 5.3.1 and Algorithm 5.4.1 is solved with an inner fixed-point iteration which is stopped as soon as an accuracy of 10^{-10} (of $\|\boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}$) is reached. Other arising linear systems are solved with GMRES (or with CG for the hybrid FEM-BEM approach) with tolerance 10^{-12}. For the spatial discretization we consider a fixed triangulation $\mathcal{T}_h$ of Ω generated by Netgen, which consists of 3939 elements and 917 nodes (prescribed mesh size $h = 1/8$).

Since the exact solution of the problem is unknown, to compute the empirical convergence rates, we consider a reference solution $\boldsymbol{m}_{h,k_{\mathrm{ref}}}$ computed with the IMEX version of the second-order midpoint scheme from [PRS18b] using the above mesh and the time-step size $k_{\mathrm{ref}} = 2 \cdot 10^{-4}$.

Figure 5.2(a) visualizes the experimental order of convergence of the four integrators. As expected, `PC2` and `PC2+IMEX` lead to second-order convergence in time. Essentially, both integrators even lead quantitatively to the same accuracy of the numerical solution. `PC1` as well as `PC1+IMEX` yield first-order convergence. Differently from the classical θ-method for linear second-order parabolic PDEs, due to the tangent plane constraint and the presence of the nodal projection, the `PC1` integrator with $\theta = 1/2$ (Crank–Nicolson-type) does not lead to any improvement of the convergence order in time (from first-order to second-order); see [AKST14] for a formal analysis in the case of the tangent plane scheme.

In Figure 5.2(b), we plot the cumulative computational costs for the integration up to the final time T. The computational effort improves considerably if the lower-order terms (i.e., the stray field) are integrated explicitly in time, since then the costly inner fixed-point iteration to solve (5.20) is omitted. Due to the more sophisticated corrector step in Algorithm 5.4.1 and Algorithm 5.4.7, the second-order schemes `PC2` and `PC2+IMEX` are (slightly) more costly than their first-order counterparts `PC1` and `PC1+IMEX`, respectively.

Further, we repeat the experiment for different values of $\theta \in [0,1]$ for both, `PC1+IMEX` and `PC2+IMEX`. The results for `PC1+IMEX` in Figure 5.3(a) confirm that the strong CFL condition $k = o(h^2)$, which is imposed to obtain stability and convergence of `PC1+IMEX` (see Remark 5.3.7(ii)) with $\theta < 1/2$, are also crucial in practice. As expected, the observed order of convergence of `PC1+IMEX` is unaffected by the choice of $\theta \in [0,1]$.

The results for `PC2+IMEX` shown in Figure 5.3(b) are quite surprising: While for $\theta \neq 1/2$, the simulation is not stable for larger time-step sizes $k > 0$, still second-order convergence is observed for all $0 \leq \theta \leq 1$ as the time-step size k decreases below a certain threshold. The preserved second-order accuracy for $\theta \neq 1/2$ might be a consequence of the degree of implicitness θ only appearing in the predictor, but not in the corrector of the scheme. In
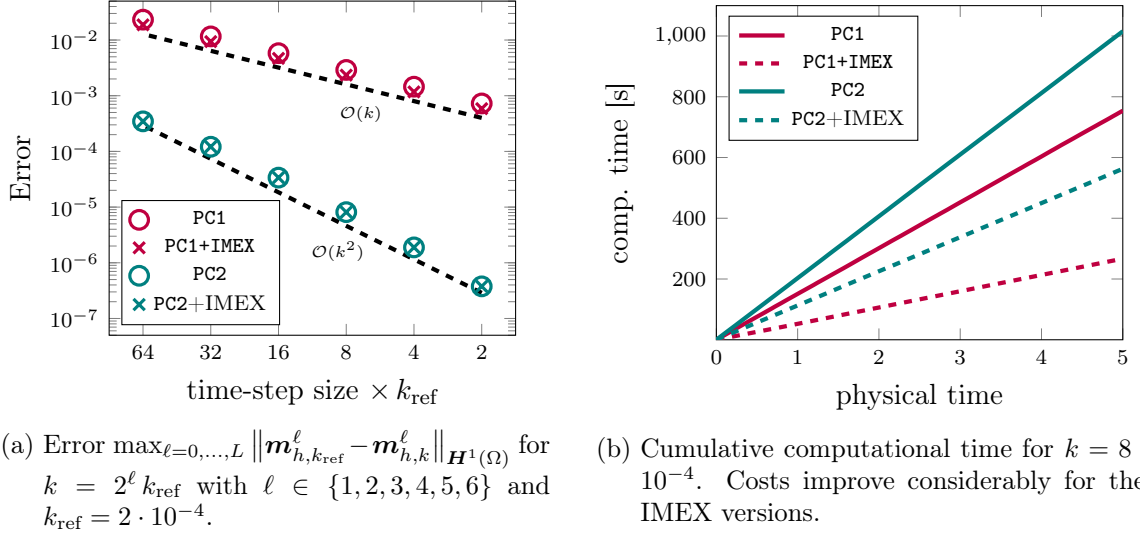
(a) Error $\max_{\ell=0,\ldots,L} \left\| \boldsymbol{m}^{\ell}_{h,k_{\mathrm{ref}}} - \boldsymbol{m}^{\ell}_{h,k} \right\|_{\boldsymbol{H}^1(\Omega)}$ for $k = 2^{\ell} k_{\mathrm{ref}}$ with $\ell \in \{1, 2, 3, 4, 5, 6\}$ and $k_{\mathrm{ref}} = 2 \cdot 10^{-4}$.

(b) Cumulative computational time for $k = 8 \cdot 10^{-4}$. Costs improve considerably for the IMEX versions.

Figure 5.2: Experiments of Section 5.5.2: Order of convergence (left) and cumulative computational time (right) of the integrators for $\theta = 1/2$.

contrast to stability for `PC1+IMEX`, the results of this experiment indicate that for stability of `PC2+IMEX` more restrictive CFL conditions are necessary for $\theta \neq 1/2$ than for $\theta = 1/2$. This observation is further investigated in Section 5.5.3.

Overall, the proposed `PC2+IMEX` integrator with $\theta = 1/2$ appears to be the method of choice with respect to experimental stability, computational time, and empirical accuracy.

5.5.3 Experimental stability of `PC2`

We demonstrated the potential of (the IMEX version of) the second-order predictor-corrector scheme `PC2` (`PC2+IMEX`) in Section 5.5.1 and Section 5.5.2. Our analysis guarantees unconditional well-posedness of the proposed second-order integrators in theory (Theorem 5.4.3) and in practice (Theorem 5.4.5). However, neither the present work nor [KW18] include a rigorous analysis on the stability of the second-order predictor-corrector scheme `PC2` (Algorithm 5.4.1), or its variant `PC2+IMEX` (Algorithm 5.4.7). More precisely, it remains unclear whether the prescription of a CFL condition $k = o(h^{\beta})$ for some $\beta > 0$ is sufficient to prove a discrete energy estimate of the form

$$\|\boldsymbol{\nabla m}^J_h\|^2 \leq \|\boldsymbol{\nabla m}^0_h\|^2 \qquad \text{for all } J = 0, \ldots, L, \tag{5.36}$$

where we omitted any lower-order contributions; see, e.g., (5.26) for the full discrete energy estimate for `PC1+IMEX`.

Hence, we close this section by a numerical study investigating the stability of `PC2`. Note that `PC2+IMEX` coincides with `PC2` for the exchange only case $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell^2_{\mathrm{ex}} \boldsymbol{\Delta m}$ of LLG, which is considered in the following experiments. Motivated by the observations on stability of `PC2+IMEX` in Figure 5.3(b), particular focus is put on the dependence on $0 \leq \theta \leq 1$, which controls the degree of implicitness in the predictor step (5.29).
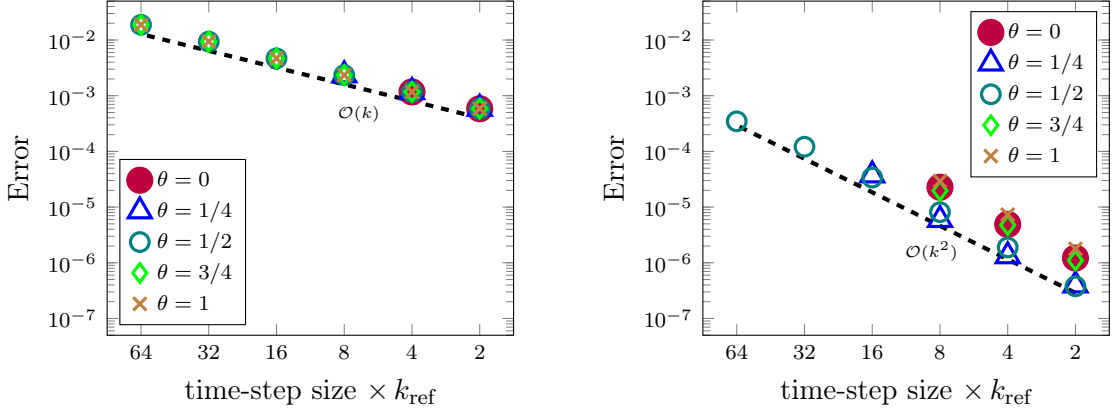
(a) Recomputation of Figure 5.2(a) for `PC1+IMEX` with various $\theta \in [0, 1]$.

(b) Recomputation of Figure 5.2(a) for `PC2+IMEX` with various $\theta \in [0, 1]$.

Figure 5.3: Experiments of Section 5.5.2: Order of convergence and stability for `PC1+IMEX` and `PC2+IMEX` for different values of $\theta \in [0, 1]$. Stability is lost for `PC1+IMEX` (left) with $\theta = 0$ for $k \geq 8 \cdot k_{\text{ref}}$, and with $\theta = 1/4$ for $k \geq 16 \cdot k_{\text{ref}}$; for `PC2+IMEX` (right) with $\theta \in \{0, 3/4, 1\}$ for $k \geq 16 \cdot k_{\text{ref}}$, and with $\theta = 1/4$ for $k \geq 32 \cdot k_{\text{ref}}$.

Setup

We consider the partition $\mathcal{T}_h$ of the unit cube from Section 5.5.2. For a non-uniform initial condition $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, we consider the exchange only case $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \boldsymbol{\Delta m}$ of LLG and relax the dynamics until the (uniform) equilibrium state is reached. Due to the absence of any lower-order contributions ($\boldsymbol{\pi} \equiv \boldsymbol{0}, \boldsymbol{f} \equiv \boldsymbol{0}$), the equilibrium state is a uniform magnetization in space, and the simulation is successfully stopped as soon as $\|\boldsymbol{\nabla m}_h^L\|^2 \leq 10^{-8}$ for some $L > 0$. If $\|\boldsymbol{\nabla m}_h^{\ell+1}\|^2 \leq \|\boldsymbol{\nabla m}_h^\ell\|^2$ for all $\ell = 0, \ldots, L - 1$, the simulation is considered to be stable for the triangulation $\mathcal{T}_h$ with fixed time-step size $k > 0$ and initial condition $\boldsymbol{m}_h^0 \in \mathcal{M}_h$. If for some $\ell \geq 0$ the energy increases, i.e., if there holds $\|\boldsymbol{\nabla m}_h^{\ell+1}\|^2 > \|\boldsymbol{\nabla m}_h^\ell\|^2$, then we abort the simulation and we consider the simulation to be unstable for this combination of $\mathcal{T}_h$, $k > 0$, and $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.

Random initial state

We choose the initial state $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ such that $\{\boldsymbol{m}_z(\boldsymbol{z})\}_{\boldsymbol{z} \in \mathcal{N}_h}$ is distributed randomly on $\mathbb{S}^2$.

Figure 5.4 shows that for any fixed $0 \leq \theta \leq 1$ the simulation is stable if the time-step size $k > 0$ is chosen small enough. Clearly, stability of the simulation does not only depend on the chosen time-step size $k > 0$, but also on the parameter θ: Values of θ close to $1/2$ (best at 0.4375 in this experiment) appear to be far less restrictive for the time-step size $k > 0$ than values farther from $1/2$. We note that we repeated this experiment for various random initial states, all producing essentially the same result (not displayed).
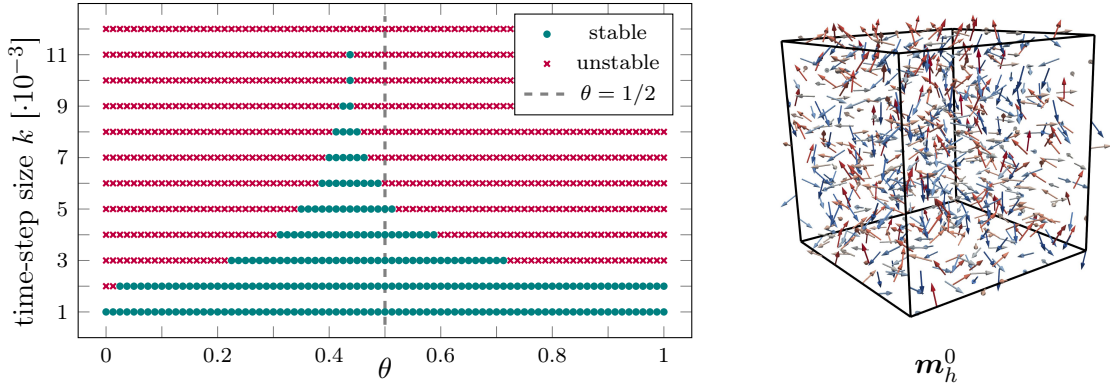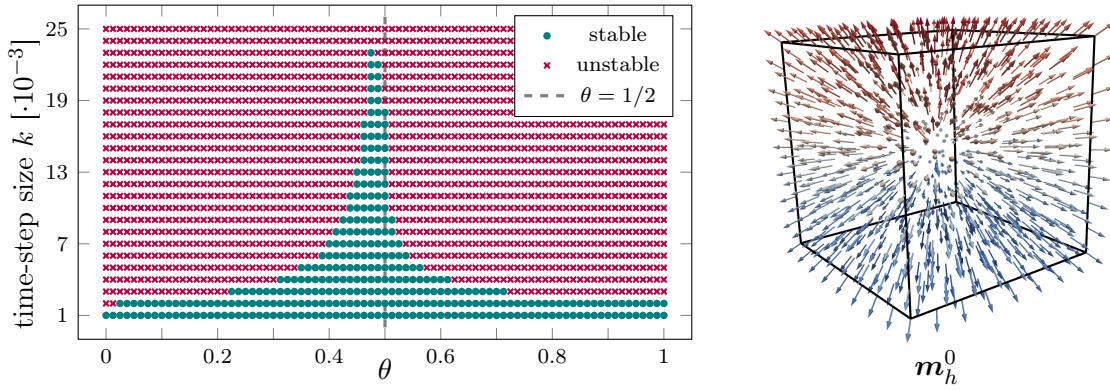
Figure 5.4: Experiment of Section 5.5.3: Right: Random state $\boldsymbol{m}_h^0$ colored by the z-component; red pointing upwards, blue downwards. Left: For all $\theta = 0/80, 1/80, \ldots, 80/80$ and all $k = 1 \cdot 10^{-3}, 2 \cdot 10^{-3}, \ldots, 12 \cdot 10^{-3}$, the stability of PC2 is investigated.

Hedgehog state

We repeat the experiment from Section 5.5.3 for $\boldsymbol{m}_h^0$ being the so-called hedgehog state, i.e., considering the cube to be centered around the origin, for each vertex $\boldsymbol{z} \in \mathcal{N}_h$ we set the initial value $\boldsymbol{m}_h^0(\boldsymbol{z}) := \boldsymbol{z}/|\boldsymbol{z}| \in \mathbb{S}^2$.

Figure 5.5: Experiment of Section 5.5.3: Right: Hedgehog state $\boldsymbol{m}_h^0$ colored by the z-component; red pointing upwards, blue downwards. Left: For all $\theta = 0/80, 1/80, \ldots, 80/80$ and all $k = 1 \cdot 10^{-3}, 2 \cdot 10^{-3}, \ldots, 25 \cdot 10^{-3}$, the stability of PC2 is investigated.

Figure 5.5 shows that again for any $0 \leq \theta \leq 1$ the simulation is stable if the time-step size $k > 0$ is chosen small enough. As in Section 5.5.3, values of θ close to $1/2$ appear to be far less restrictive for the time-step size $k > 0$ than values farther from $1/2$, with the optimal choice this time closer to $1/2$, precisely at $\theta = 0.475$. Interestingly, for the parameter $\theta \in [0, 1]$ chosen far from $1/2$, the results quantitatively match with those for the random initial state from Section 5.5.3. Closer to $1/2$, however, much larger time-step

sizes $k > 0$ allow for stable simulations as for the random initial state.

Variation of the Gilbert damping parameter

We repeat the experiment from Section 5.5.3 for different values of $\alpha = 1/2, 1/4, 1/8, 1/16$.
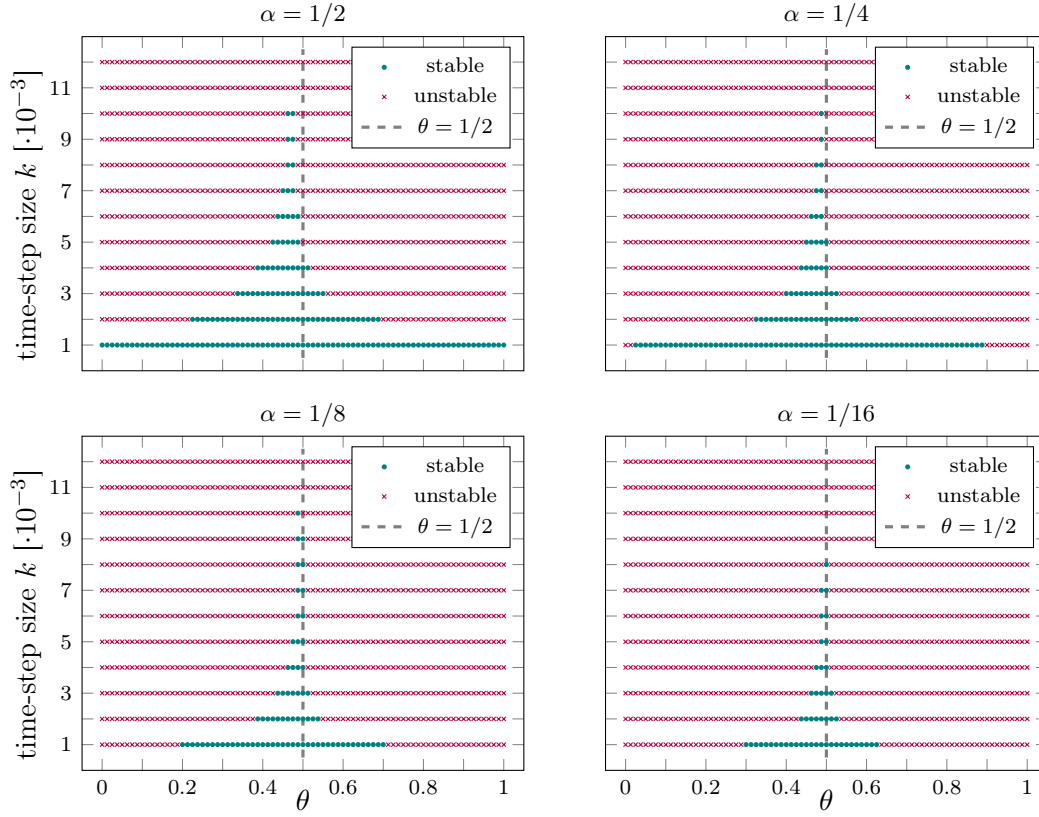
Figure 5.6: Experiment of Section 5.5.3: With $\boldsymbol{m}_h^0$ the random state from Figure 5.4(right) and different damping parameters $\alpha = 1/2, 1/4, 1/8, 1/16$, for all $\theta = 0/80, 1/80, \ldots, 80/80$ and all $k = 1 \cdot 10^{-3}, 2 \cdot 10^{-3}, \ldots, 12 \cdot 10^{-3}$, the stability of `PC2(+IMEX)` is investigated.

Figure 5.6 shows that, if the damping parameter α decreases, smaller time-step sizes $k > 0$ are necessary to obtain stable simulations with `PC2(+IMEX)`. This observation is in agreement with the role played by α in the model, i.e., incorporating dissipation. Again, as previously observed for $\alpha = 1$, values of θ close to $1/2$ allow for larger time-step sizes $k > 0$ than values farther from $1/2$; with the least restrictive choices at $\theta = 0.4375$ for $\alpha = 1$, $\theta = 0.4625$ to 0.475 for $\alpha = 1/2$, $\theta = 0.4875$ for $\alpha = 1/4$, $\theta = 0.4875$ for $\theta = 1/8$, and $\theta = 0.5$ for $\alpha = 1/16$. We obtain analogous results when varying α for the initial hedgehog state (not displayed).

Concluding remarks on the stability of the second-order scheme

All experiments in this section show that, in contrast to `PC1` (Theorem 5.3.6), larger values of θ do not improve stability of the second-order scheme `PC2`. On the contrary, it is even the case that large values of θ perform as bad as small values of θ. For a generic simulation with `PC2(+IMEX)`, we suggest to pick the degree of implicitness $\theta = 1/2$ in the predictor. Although, when considering one particular simulation setup, there might be better choices allowing for even larger time-step sizes, the choice $\theta = 1/2$ performed reliably throughout all experiments. In particular, the results from Section 5.5.3 indicate that the deterioration of the optimal θ (with respect to stability) from $1/2$ might occur specifically for large values of α, and quickly vanish as the damping parameter α decreases. Moreover in future works, proving stability of `PC2` under some CFL condition for the special case $\theta = 1/2$ might be a possible first step in theoretically understanding stability of `PC2`. This seems reasonable, as in this special case only the same highest-order term $\boldsymbol{\Delta}_h(\boldsymbol{m}_h^\ell + (k/2)\boldsymbol{v}_h^\ell)$ appears in the predictor and the corrector of `PC2`. Hence these terms partially cancel out, when subtracting the two equations (5.29)–(5.30) from each other.

6 Computational micromagnetics with Commics

This chapter consists of the article [PRS$^+$20] together with Michele Ruggeri, Bernhard Stiftner, Lukas Exl, Matthias Hochsteger, Gino Hrkac, Joachim Schöberl, Norbert J. Mauser, and Dirk Praetorius. Additionally, Section 6.5 contains various simulation scripts from [PRS$^+$18a], which were excluded in the article [PRS$^+$20] due to space considerations.

Abstract. We present our open-source Python module *Commics* for the study of the magnetization dynamics in ferromagnetic materials via micromagnetic simulations. It implements state-of-the-art unconditionally convergent finite element methods for the numerical integration of the Landau–Lifshitz–Gilbert equation. The implementation is based on the multiphysics finite element software *Netgen/NGSolve*. The simulation scripts are written in Python, which leads to very readable code and direct access to extensive post-processing. Together with documentation and example scripts, the code is freely available on GitLab.

6.1 Introduction

Micromagnetism is a continuum theory for ferromagnetic materials located between quantum theory and Maxwell's electromagnetism [Bro63, Aha01, Kro07]. The magnetization distribution is modeled as a continuous vector field, where nonlocal magnetostatic interactions and local contributions are taken into account. Typical micromagnetic models exhibit length scales ranging from nanometers to few micrometers, which is often infeasible for atomistic spin dynamics simulations. In recent decades, micromagnetics evolved as a computational field, that nowadays represents a successful tool for numerical studies in materials science with important contemporary applications, e.g., data storage structures like hard disk drives [SVA$^+$15, KOSS16], random access memories [MSOS12], or nanowires [SZZS00, Her01], magnetologic devices [BSH$^+$08], soft magnetic sensor systems [BSP$^+$17], and high performance permanent magnets [SAON$^+$13, BOS$^+$14, FKO$^+$17].

6.1.1 Existing software

In recent years, advances in computer architecture, programming environments, and numerical methods led to the development of several micromagnetic codes, which aim at the numerical integration of the fundamental equation in micromagnetics, the Landau–Lifshitz–Gilbert equation (LLG), a well-accepted model for the magnetization dynamics [LL08, Gil55]. Well-known simulation packages based on finite difference discretizations [MD07] on Cartesian grids are *OOMMF* [DP99], recently extended for GPU usage [FCH$^+$16] and

endowed with a user-friendly Python interface [BPF17], *MuMax*^{3} (GPU) [VLD$^+$14], *MicroMagnum* [mic] (CPU and GPU), *magnum.fd* [mag], and *Fidimag* [BCOnP$^+$18]. These implementations are both memory and computationally efficient owing to the uniform mesh and the particularly advantageous utilization of the fast Fourier transform for the long range field part [BRH93, AES$^+$13]. However, approaches based on the Finite Element Method (FEM) are geometrically more flexible [SHB$^+$07] and provided in the scientific codes *MagPar* [SFS$^+$03], *TetraMag* [KWH10] and its successor *tetmag2*, *Nmag* [FFBF07] and its successor *Finmag* [BBW$^+$], as well as in commercial software like *FastMag* [CLL$^+$11], *FEMME* [fem], and *magnum.fe* [AEB$^+$13].

6.1.2 Numerical analysis

For an introduction to the mathematical analysis of numerical integrators for dynamic micromagnetic simulations, we refer to the monographs [Pro01, BBNP14], the review articles [KP06, GC07, Cim08], and the references therein. The ultimate goal is the development of unconditionally convergent integrators, i.e., numerical schemes for which (a subsequence of) the output converges towards a weak solution of LLG in the sense of [AS92] without requiring any restrictive CFL-type coupling condition on the temporal and spatial discretization parameters. In our work, we consider two types of methods characterized by such good theoretical properties: the *tangent plane scheme* [Alo08, AKST14] and the *midpoint scheme* [BP06]; see Section 6.3 for more details.

6.1.3 Contributions

This work presents our novel open-source Python module *Commics* (COmputational MicroMagnetICS) to perform computational studies of the magnetization dynamics in ferromagnetic materials via micromagnetic simulations. The software is based on the multiphysics finite element software *Netgen/NGSolve* [ngs] and is made user-friendly by a high-level Python interface. While many existing codes popular in the physics community are fairly performance optimized, they often lack a thorough mathematical convergence analysis. In contrast to that, our implementation arises from recent results in the numerical analysis of unconditionally convergent LLG integrators. The code is freely available on GitLab [Pfe] together with documentation and example scripts.

6.1.4 Outline

This work is organized as follows: We fix the notation and the precise micromagnetic setting in Section 6.2. In Section 6.3, the implemented algorithms are briefly presented. Section 6.4 demonstrates the exemplary use of *Netgen/NGSolve* (*NGS*) and discusses the integration of the boundary element library *BEM++* [ŚBA$^+$15]. Finally, Section 6.5 provides Python scripts for several benchmark problems, in order to verify our module and to demonstrate its usage.

6.2 Micromagnetic setting

Let $\Omega \subset \mathbb{R}^3$ denote the volume occupied by a ferromagnet. In micromagnetics, the quantity of interest is the magnetization $\boldsymbol{M} \colon \Omega \to \mathbb{R}^3$ (in A/m). If the temperature is constant and far below the so-called Curie temperature of the material, the modulus of the magnetization is constant, i.e., it holds that $|\boldsymbol{M}| = M_{\mathrm{s}}$ with $M_{\mathrm{s}} > 0$ being the saturation magnetization (in A/m). Let $\boldsymbol{m} := \boldsymbol{M}/M_{\mathrm{s}}$ denote the normalized magnetization. The magnetic state of Ω is described in terms of the magnetic Gibbs free energy (in J)

$$
\begin{aligned}
\mathcal{E}(\boldsymbol{m}) = {} & A \int_{\Omega} |\nabla \boldsymbol{m}|^2 \, dx \, + \, K \int_{\Omega} 1 - (\boldsymbol{a} \cdot \boldsymbol{m})^2 \, dx \, + \, D \int_{\Omega} (\nabla \times \boldsymbol{m}) \cdot \boldsymbol{m} \, dx \\
& - \frac{\mu_0 M_{\mathrm{s}}}{2} \int_{\Omega} \boldsymbol{H}_{\mathrm{s}} \cdot \boldsymbol{m} \, dx \, - \, \mu_0 M_{\mathrm{s}} \int_{\Omega} \boldsymbol{H}_{\mathrm{ext}} \cdot \boldsymbol{m} \, dx.
\end{aligned}
\tag{6.1}
$$

The energy in (6.1) is the sum of exchange energy, uniaxial anisotropy, bulk Dzyaloshinskii–Moriya interaction (DMI), magnetostatic energy, and Zeeman contribution, respectively. The involved material parameters and physical constants are the exchange stiffness constant $A > 0$ (in J/m), the anisotropy constant $K \geq 0$ (in J/m³), the easy axis $\boldsymbol{a} \in \mathbb{R}^3$ with $|\boldsymbol{a}| = 1$ (dimensionless), the DMI constant $D \in \mathbb{R}$ (in J/m²), and the vacuum permeability $\mu_0 = 4\pi \cdot 10^{-7} \, \mathrm{N/A}^2$. Moreover, $\boldsymbol{H}_{\mathrm{ext}}$ and $\boldsymbol{H}_{\mathrm{s}}$ denote the applied external field (assumed to be unaffected by variations of $\boldsymbol{m}$) and the stray field, respectively (both in A/m). The stray field (sometimes also referred to as demagnetizing or dipolar field) solves the magnetostatic Maxwell equations

$$
\nabla \cdot (\boldsymbol{H}_{\mathrm{s}} + M_{\mathrm{s}} \boldsymbol{m} \chi_{\Omega}) = 0 \quad \text{in } \mathbb{R}^3,
\tag{6.2a}
$$

$$
\nabla \times \boldsymbol{H}_{\mathrm{s}} = \boldsymbol{0} \quad \text{in } \mathbb{R}^3,
\tag{6.2b}
$$

where $(\boldsymbol{m} \chi_{\Omega})(\boldsymbol{x}) = \boldsymbol{m}(\boldsymbol{x})$ in Ω and $(\boldsymbol{m} \chi_{\Omega})(\boldsymbol{x}) = 0$ elsewhere. Stable magnetization configurations are those which minimize the magnetic Gibbs free energy (6.1). The dynamics towards equilibrium of the magnetization is governed by LLG

$$
\partial_t \boldsymbol{m} = -\gamma_0 \, \boldsymbol{m} \times \left[\boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m}) + \boldsymbol{T}(\boldsymbol{m}) \right] + \alpha \, \boldsymbol{m} \times \partial_t \boldsymbol{m} \qquad \text{in } (0, \infty) \times \Omega,
\tag{6.3a}
$$

$$
\partial_{\boldsymbol{n}} \boldsymbol{m} = -\frac{D}{2A} \, \boldsymbol{m} \times \boldsymbol{n} \qquad \text{on } (0, \infty) \times \partial\Omega,
\tag{6.3b}
$$

$$
\boldsymbol{m}(0) = \boldsymbol{m}^0 \quad \text{with } |\boldsymbol{m}^0| = 1 \qquad \text{in } \Omega.
\tag{6.3c}
$$

In (6.3), $\gamma_0 = 2.212 \cdot 10^5 \, \mathrm{m/(A \, s)}$ is the gyromagnetic ratio of the electron, $\alpha \in (0, 1]$ is the dimensionless Gilbert damping parameter, and $\boldsymbol{n} \colon \partial\Omega \to \mathbb{R}^3$ with $|\boldsymbol{n}| = 1$ denotes the outward-pointing unit normal vector to $\partial\Omega$. The effective field $\boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m})$ is related to the functional derivative of the energy with respect to the magnetization and takes the form

$$
\begin{aligned}
\boldsymbol{H}_{\mathrm{eff}}(\boldsymbol{m}) := {} & -\frac{1}{\mu_0 M_{\mathrm{s}}} \frac{\delta \mathcal{E}(\boldsymbol{m})}{\delta \boldsymbol{m}} \\
= {} & \frac{2A}{\mu_0 M_{\mathrm{s}}} \Delta \boldsymbol{m} + \frac{2K}{\mu_0 M_{\mathrm{s}}} (\boldsymbol{a} \cdot \boldsymbol{m}) \boldsymbol{a} - \frac{2D}{\mu_0 M_{\mathrm{s}}} \nabla \times \boldsymbol{m} + \boldsymbol{H}_{\mathrm{s}} + \boldsymbol{H}_{\mathrm{ext}}.
\end{aligned}
\tag{6.4}
$$

Finally, the term $\boldsymbol{T(m)}$ collects all nonenergetic torque terms, which arise, e.g., when an electric current flows in a conducting ferromagnet. For instance, the Oersted field $\boldsymbol{T(m)} = \boldsymbol{H}_{\mathrm{c}}$ (in A/m) is described by the magnetostatic Maxwell equations

$$\nabla \cdot \boldsymbol{H}_{\mathrm{c}} = 0 \qquad \text{in } \mathbb{R}^3, \tag{6.5a}$$

$$\nabla \times \boldsymbol{H}_{\mathrm{c}} = \boldsymbol{J}_{\mathrm{e}} \chi_\Omega \quad \text{in } \mathbb{R}^3, \tag{6.5b}$$

where $\boldsymbol{J}_{\mathrm{e}}$ denotes the electric current density (in A/m^2). Two other prominent examples are related to the so-called spin transfer torque [Slo96, Ber96], which arises in the presence of spin-polarized currents. The Slonczewski contribution [Slo96], which takes the form

$$\boldsymbol{T(m)} = \frac{\hbar |\boldsymbol{J}_{\mathrm{e}}| G(\boldsymbol{m} \cdot \boldsymbol{p}, P)}{e\mu_0 M_{\mathrm{s}} d} \boldsymbol{m} \times \boldsymbol{p},$$

$$\text{with} \quad G(\boldsymbol{m} \cdot \boldsymbol{p}, P) = \left[\frac{(1+P)^3(3 + \boldsymbol{m} \cdot \boldsymbol{p})}{4P^{3/2}} - 4 \right]^{-1}, \tag{6.6}$$

is used for the simulation of switching processes in structures with current-perpendicular-to-plane injection geometries, e.g., magnetic multilayers. The involved physical quantities are the reduced Planck constant $\hbar > 0$ (in J s), the elementary charge $e > 0$ (in A s), a dimensionless polarization parameter $P \in (0,1)$, a dimensionless unit vector $\boldsymbol{p} \in \mathbb{R}^3$ representing the magnetization of a uniformly-magnetized polarizing layer (the so-called fixed layer), and the thickness $d > 0$ of the so-called free layer (in m). The Zhang–Li contribution [ZL04, TNMS05] is used for the simulation of the current-driven motion of domain walls in single-phase samples characterized by current-in-plane injection geometries and, according to [ZL04] (resp., [TNMS05]), takes the form

$$\boldsymbol{T(m)} = -\frac{1}{\gamma_0} [\boldsymbol{m} \times (\boldsymbol{u} \cdot \nabla)\boldsymbol{m} + \xi(\boldsymbol{u} \cdot \nabla)\boldsymbol{m}], \tag{6.7a}$$

$$\text{with} \quad \boldsymbol{u} = -\frac{P\mu_B}{eM_{\mathrm{s}}(1+\xi^2)} \boldsymbol{J}_{\mathrm{e}} \quad \left(\text{resp., } \boldsymbol{u} = -\frac{Pg_{\mathrm{e}}\mu_B}{2eM_{\mathrm{s}}} \boldsymbol{J}_{\mathrm{e}} \right). \tag{6.7b}$$

The involved physical quantities are the spin velocity vector $\boldsymbol{u} \in \mathbb{R}^3$ (in m/s), the dimensionless ratio of nonadiabaticity $\xi > 0$, the Bohr magneton $\mu_B > 0$ (in A m^2), and the dimensionless g-factor of the electron $g_{\mathrm{e}} \approx 2$.

In (6.1) and (6.3), to fix the ideas, we considered the case of the bulk DMI as a prototype for chiral interactions [Dzy58, Mor60]. However, our implementation also covers the case of the interfacial DMI for which, considering a thin magnetic film aligned with the $x_1 x_2$-plane, the bulk DMI contributions to the energy (6.1), to the effective field (6.4), and to the boundary conditions (6.3b) are replaced by the corresponding contributions of the interfacial DMI, which read

$$D \int_\Omega [m_3(\partial_1 m_1 + \partial_2 m_2) - (m_1\,\partial_1 m_3 + m_2\,\partial_2 m_3)]\,dx,$$

$$-\frac{2D}{\mu_0 M_{\mathrm{s}}} \begin{pmatrix} -\partial_1 m_3 \\ -\partial_2 m_3 \\ \partial_1 m_1 + \partial_2 m_2 \end{pmatrix}, \quad \text{and} \quad -\frac{D}{2A}(\boldsymbol{e}_3 \times \boldsymbol{n}) \times \boldsymbol{m},$$

respectively; see, e.g., [CL98, SCR$^+$13, HPP$^+$19]. If the DMI constant D in (6.1) equals zero (no chiral interaction), then the boundary conditions (6.3b) become homogeneous Neumann boundary conditions.

6.3 Algorithms

The algorithms implemented in our Python module *Commics* employ a uniform partition of the time interval with constant time-step size $\Delta t > 0$. For the spatial discretization, we consider a tetrahedral mesh $\mathcal{T}_h$ of the ferromagnet Ω with mesh size $h > 0$. The associated FEM space of piecewise affine and globally continuous functions reads

$$V_h := \{\varphi_h \colon \Omega \to \mathbb{R} \text{ continuous} \colon \varphi_h|_K \text{ is affine for all elements } K \in \mathcal{T}_h\}. \tag{6.8}$$

For each time-step $n = 0, 1, 2, \ldots$, we seek for approximations

$$(V_h)^3 \ni \boldsymbol{m}_h^n \approx \boldsymbol{m}(n\Delta t) \quad \text{such that} \quad |\boldsymbol{m}_h^n(\boldsymbol{z})| = 1 \text{ for all nodes } \boldsymbol{z} \text{ of } \mathcal{T}_h,$$

i.e., for any time-step, the approximate magnetization satisfies the unit-length constraint $|\boldsymbol{m}| = 1$ at the nodes of the mesh.

6.3.1 Tangent plane scheme

Tangent plane schemes (sometimes also referred to as *projection methods*) are based on variational formulations of the equivalent form of LLG

$$\alpha\, \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = \gamma_0(\boldsymbol{H}_{\text{eff}}(\boldsymbol{m}) + \boldsymbol{T}(\boldsymbol{m})) - \gamma_0[(\boldsymbol{H}_{\text{eff}}(\boldsymbol{m}) + \boldsymbol{T}(\boldsymbol{m})) \cdot \boldsymbol{m}]\boldsymbol{m}.$$

The orthogonality $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$, which characterizes any solution of (6.3a), is enforced at the discrete level by considering the discrete tangent space

$$\mathcal{K}_h(\boldsymbol{m}_h^n) := \{\varphi_h \in (V_h)^3 \colon \varphi_h(\boldsymbol{z}) \cdot \boldsymbol{m}_h^n(\boldsymbol{z}) = 0 \text{ for all nodes } \boldsymbol{z} \text{ of } \mathcal{T}_h\} \subset (V_h)^3.$$

For each time-step, one has to solve a constrained linear system to compute $\boldsymbol{v}_h^n \approx \partial_t \boldsymbol{m}(n\Delta t)$ in $\mathcal{K}_h(\boldsymbol{m}_h^n)$. With $\boldsymbol{m}_h^n$ and $\boldsymbol{v}_h^n$ at hand, one then computes

$$\boldsymbol{m}_h^{n+1} \in (V_h)^3 \quad \text{by} \quad \boldsymbol{m}_h^{n+1}(\boldsymbol{z}) := \frac{\boldsymbol{m}_h^n(\boldsymbol{z}) + \Delta t\, \boldsymbol{v}_h^n(\boldsymbol{z})}{|\boldsymbol{m}_h^n(\boldsymbol{z}) + \Delta t\, \boldsymbol{v}_h^n(\boldsymbol{z})|} \quad \text{for all nodes } \boldsymbol{z} \text{ of } \mathcal{T}_h.$$

The original tangent plane scheme from [Alo08] is formally first-order in time and was analyzed for the energy being only the exchange contribution. The scheme was extended to general lower-order effective field contributions [AKT12, BFF$^+$14], DMI [HPP$^+$19], and the coupling with other partial differential equations, e.g., various forms of Maxwell's equations [LT13, BPP15, LPPT15, FT17a], spin diffusion [AHP$^+$14], and magnetostriction [BPPR14]. A projection-free version of the method was analyzed in [AHP$^+$14, Rug16]. In a variant from [AKST14, DPP$^+$20], the formal convergence order in time has been increased from one to two. Effective solution strategies and preconditioning for the resulting constrained linear system have recently been proposed in [KPP$^+$19].

6.3.2 Midpoint scheme

The midpoint scheme is based on a variational formulation of the Gilbert form (6.3a) of LLG. It consists of two fundamental ingredients: the implicit midpoint rule in time and the mass-lumped L^2-product in space, defined as

$$\langle \boldsymbol{\varphi}, \boldsymbol{\psi} \rangle_h := \int_\Omega \mathcal{I}_h(\boldsymbol{\varphi} \cdot \boldsymbol{\psi}) \, dx \approx \langle \boldsymbol{\varphi}, \boldsymbol{\psi} \rangle_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{\varphi}, \boldsymbol{\psi} \colon \Omega \to \mathbb{R}^3 \text{ continuous.} \qquad (6.9)$$

Here, $\mathcal{I}_h$ is the standard nodal interpolant associated with $\mathcal{T}_h$. The resulting scheme is second-order in time, inherently preserves the unit-length constraint and the energy of the solutions, but requires the solution of one *nonlinear* system for $\boldsymbol{m}_h^{n+1/2} := (\boldsymbol{m}_h^{n+1} + \boldsymbol{m}_h^n)/2 \in (V_h)^3$ per time-step. The midpoint scheme was proposed and analyzed in [BP06]. The scheme was extended to lower-order terms [PRS18b], the coupling with the Maxwell equation [BBP08], and a variant of LLG in heat-assisted magnetic recording [BPS09, BPS12]. The resulting nonlinear system is usually solved with constraint preserving fixed-point iterations, which, however, spoil the unconditional convergence; see [BP06, Section 4] or [PRS18b, Section 5]. For the linear system arising in each iteration of the fixed-point solver, we employ diagonally preconditioned GMRES [SS86]. For another approach, we refer to, e.g., [Baň10].

6.3.3 Magnetostatic Maxwell equations

For the computation of stray field and Oersted field, i.e., for the numerical solution of the magnetostatic Maxwell equations (6.2) and (6.5), we follow a hybrid FEM-BEM approach, which combines FEM with the boundary element method (BEM); see [FK90, HK14]. The method uses the superposition principle and computes the magnetic scalar potential u such that $\boldsymbol{H}_\mathrm{s} = -\nabla u$ and the Oersted field $\boldsymbol{H}_\mathrm{c}$ by splitting the problem into two parts, where BEM techniques for the evaluation of the double-layer potential are employed. For the explicit algorithms implemented in *Commics*, see [PRS18b, Chapter 4.1] for the stray field and [Rug16, Section 5.2.4.3] for the Oersted field, respectively.

6.4 Implementation

Our Python module *Commics* is based on the *Netgen/NGSolve* [ngs] (*NGS*) FEM software and provides a tool to perform micromagnetic simulations with the algorithms described in Section 6.3 for a variety of energy contributions and dissipative effects. It is purely Python-based, provides extensive simulation data for reproducibility and post-processing, and automatically takes care of, e.g., the definition and the assembly of underlying (bi)linear forms as well as the numerical solution of the arising linear and nonlinear systems. Although not needed for the use of *Commics*, this section advertises some of the core features of *NGS*, as well as the coupling of *NGS* with *BEM++*.

6.4.1 Basic features

Geometry handling, mesh-generation, FEM spaces, and assembly routines are intentionally hidden from the user of *Commics* and are internally covered in the *NGS* framework. For

instance, given an *NGS* object `mesh`, representing a tetrahedral mesh of the domain Ω, the discrete vector-valued product space $(V_h)^3$ is simply generated by

```
Vh3 = VectorH1(mesh, order=1)
```

where the syntax `VectorH1` indicates that `Vh3` is a proper subspace of the vector-valued Sobolev space $(H^1(\Omega))^3$.

6.4.2 Symbolic (bi)linear forms

Although the use of *Commics* requires minimal knowledge of *NGS*, we shortly describe one feature of the library, namely the definition of (bi)linear forms: *NGS* allows the symbolic definition of (time-dependent) (bi)linear forms. For instance, given the space `Vh3` defined in Section 6.4.1, the LLG-specific cross-product bilinear form

$$\langle \boldsymbol{m}_h^n \times \boldsymbol{\psi}_h, \boldsymbol{\varphi}_h \rangle_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{\varphi}_h, \boldsymbol{\psi}_h \in (V_h)^3 \tag{6.10}$$

is symbolically defined on the Python level by the following code snippet:

```
# grid-, trial- and testfunction
psi = Vh3.TrialFunction()
phi = Vh3.TestFunction()
m = GridFunction(Vh3)
# bilinear form
LHS = BilinearForm(Vh3)
ir = IntegrationRule(TET, order=3)
LHS += SymbolicBFI(Cross(m, psi) * phi, intrule=ir)
```

In the last line, `SymbolicBFI` with the test function `phi` and the trial function `psi` realizes the bilinear form (6.10) and adds it to the left-hand side `LHS` for further handling. Note that via the `intrule` option, the order of the quadrature is explicitly set to 3, since the realization of the bilinear form corresponds to the exact integration of piecewise cubic polynomials on tetrahedra, hence also the parameter `TET`. For the midpoint scheme, the corresponding bilinear form employs the mass-lumped L^2-product from (6.9), i.e.,

$$\langle \boldsymbol{m}_h^n \times \boldsymbol{\psi}_h, \boldsymbol{\varphi}_h \rangle_h \quad \text{for all } \boldsymbol{\varphi}_h, \boldsymbol{\psi}_h \in (V_h)^3. \tag{6.11}$$

This bilinear form can be defined in the same way as above specifying the corresponding quadrature rule on the reference tetrahedron $\text{conv}\{\boldsymbol{0}, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$:

```
ir = IntegrationRule(points=[(0,0,0), (1,0,0), (0,1,0), (0,0,1)], \
                     weights=[1/24, 1/24, 1/24, 1/24])
```

6.4.3 Coupling with *BEM++*

For the approximate computation of the magnetostatic fields with the hybrid FEM-BEM approach from Section 6.3.3, the evaluation of the double-layer potential is required. To that end, we incorporate into *Commics* the corresponding functionality of the BEM software *BEM++* [ŚBA+15], including matrix compression techniques [Hac99]. The coupling of *NGS* and *BEM++* is done on the Python level with the *ngbem* module [Rie]; see Section 6.5.8.

6.5 Using Commics: Standard problems and numerical experiments

In this section, we present some numerical experiments performed with *Commics*. For each proposed example, we also include the executable Python script to run the simulation.

6.5.1 Using *Commics*

To run a micromagnetic simulation with *Commics*, the user has to define an object of class `commics.Integrator` and call its `Integrate` method.

The essential inputs to initialize such an object are the geometry as a `commics.Geometry` object defining the ferromagnetic domain and the meshing strategy (see Section 6.5.2), as well as an object of type `commics.Parameters` specifying, among other things, material parameters, an applied field or current, the initial magnetization state, and the time discretization strategy. Moreover, the desired time integration scheme has to be chosen. The algorithms for the numerical integration of LLG implemented in *Commics* described in Section 6.3 can be selected in the following way:

- The first-order tangent plane scheme from [Alo08] with explicit integration of the lower-order terms in time [AKT12, BFF+14] is provided as `TPS1`.

- The projection-free tangent plane scheme (with explicit integration of the lower-order terms in time) from [AHP+14, Rug16] is provided as `TPS1PF`.

- The second-order tangent plane scheme from [AKST14] and its improved version from [DPP+20] are available as `TPS2` and `TPS2AB`, respectively.

- The midpoint scheme from [BP06] and its improved version from [PRS18b] are provided as `MPS` and `MPSAB`, respectively.

Further details can be accessed with the Python built-in `help()` function available for *Commics* classes.

6.5.2 Geometry specification and meshing

Commics provides several ways to specify a geometry and to generate a corresponding mesh:

- For complex geometries, the submodule `netgen.csg` of *NGS* provides a rich number of possibilities to define geometries; see [ngs].

- For geometries often encountered in micromagnetics, e.g., cuboids and disks, one can simply provide the dimensions of the sample, a scale factor, and the desired maximum mesh size; see, e.g., the code snippets in Section 6.5.3 and Section 6.5.6. Then, `netgen.csg` will automatically be used with appropriate settings.

- Existing *NGS* meshes/geometries (stored as `.vol`-files) can simply be loaded; see, e.g., the code snippet in Section 6.5.3.
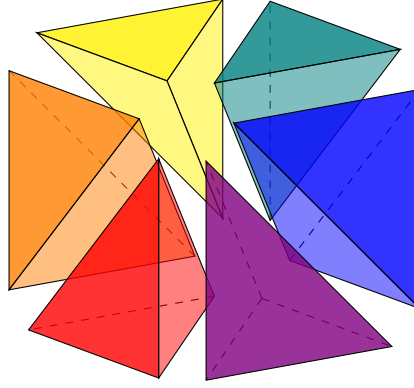
Figure 6.1: Uniform decomposition of a cube into six tetrahedra.

Meshes automatically generated by *NGS* are unstructured and obtained by the advancing front method; see [Sch97] for details. However, due to mesh quality and shape optimizations, these meshes do not necessarily satisfy the prescribed maximum mesh size. *Commics* provides two possibilities to bypass this drawback: For cuboidal geometries, setting the *Commics* option `structuredMesh=True` allows for structured meshes. First, the sample is uniformly split into cuboidal cells of prescribed size. Then, each cell is split into six tetrahedra in such a way that any tetrahedron has three mutually perpendicular edges; see Figure 6.1. This strategy is for example used in Section 6.5.4. For general geometries, *Commics* provides the means to repeatedly generate a new mesh (by prescribing a smaller and smaller mesh size each time) in *NGS*, until the initial mesh size specification is satisfied. This strategy can be enabled by setting `forceNetgenMeshSize=True` as done in Section 6.5.6.

6.5.3 μMAG standard problem #4

To describe the key aspects of a *Commics* script, we consider the μMAG standard problem #4 [MUM].

The objective is the simulation of the magnetization dynamics in a thin permalloy film of dimensions $500\,\text{nm} \times 125\,\text{nm} \times 3\,\text{nm}$ under the influence of a constant applied external field. We split the experiment into two parts: In the first stage, we obtain the so-called equilibrium S-state, which is saved to serve as the initial configuration for the second stage, where the switching dynamics is simulated.

Obtaining the S-state

We consider a structured tetrahedral mesh of the given cuboid into cells of size $h_x \times h_y \times h_z$, which are decomposed into tetrahedra as depicted in Figure 6.1. The dimension of the cells is chosen as $h_x = h_y = 125/69\,\text{nm}$ and $h_z = 1.5\,\text{nm}$. This corresponds to $228\,528$ elements with diameter $h = 2.97\,\text{nm}$, $58\,170$ vertices, as well as $78\,936$ surface elements. The material parameters of permalloy read $M_s = 8 \cdot 10^5\,\text{A/m}$, $A = 1.3 \cdot 10^{-11}\,\text{J/m}$, $D = 0\,\text{J/m}^2$, and $K = 0\,\text{J/m}^3$. To speed up the process, we deliberately choose the large value $\alpha = 1$

for the Gilbert damping parameter. For the simulation, we use a constant time-step size $\Delta t = 0.1 \, \text{ps}$.

The problem description suggests to obtain the S-state by applying a slowly reducing external field pointing in the $(1, 1, 1)$-direction. We start with a uniform initial state $\boldsymbol{m}_h^0 \equiv (1, 0, 0)$ and let the magnitude $|\boldsymbol{H}_{\text{ext}}|$ of the external field decrease linearly from $30/\mu_0$ to $0 \, \text{mT}$ over a period of $1 \, \text{ns}$. In *Commics* scripts, non-constant fields can be described using time- and space-dependent Python `lambda`-functions. Further, we relax the system for $1 \, \text{ns}$ without applying any external field and store the obtained S-state as `sp4sState.vtk` for later use.

```python
from commics import *
# specify geometry and parameters
h_xy, h_z = 125/69, 1.5
geo = Geometry(geometry=Cuboid(500, 125, 3), meshSize=(h_xy, h_xy, h_z), \
               structuredMesh=True, scaling=1e-9)
par = Parameters(A=1.3e-11, Ms=8e+5, K=0, D=0, gamma0=2.211e+5, \
                 alpha=1.0, m0=(1.0, 0.0, 0.0), \
                 T_start=0, T_end=1e-9, timeStepSize=0.1e-12)
# define time dependent applied field via lambda function
from math import sqrt
field = lambda t,x,y,z : (par.T_end - t) / (par.T_end - par.T_start) \
        * 30.0e-3 / par.mu0 / sqrt(3.0)
par.H_ext = (field, field, field)
# define integrator and run simulation from T_start to T_end
sp4 = Integrator(scheme=TPS2AB, geometry=geo, parameters=par)
sp4.Integrate()
# Relax for another nanosecond and save the mesh and the result
sp4.Integrate(duration=1e-9, relax=True)
sp4.SaveMesh("sp4mesh")
sp4.SaveMagnetization("sp4sState")
```

Switching

We assume that the folder `data` contains the two files `sp4sState.vtk` and `sp4mesh.vol` saved from the simulation described in Section 6.5.3. As stated in the problem description, we choose $\alpha = 0.02$ and set the external field to $\boldsymbol{H}_{\text{ext}} = (-24.6, 4.3, 0)/\mu_0 \, \text{mT}$. Then, we run the simulation for $3 \, \text{ns}$.

```python
from commics import *
# specify geometry and parameters
geo = Geometry(geometry="data/sp4mesh.vol", scaling=1e-9)
par = Parameters(A=1.3e-11, Ms=8e+5, K=0, D=0, gamma0=2.211e+05, \
                 alpha=0.02, timeStepSize=0.1e-12, m0="data/sp4sState.vtk")
par.H_ext = (-24.6e-3/par.mu0, 4.3e-3/par.mu0, 0.0)
# define integrator and run simulation
sp4 = Integrator(scheme=TPS2AB, geometry=geo, parameters=par)
sp4.Integrate(duration=3e-9)
```

For comparison, the desired output of this benchmark problem is the evolution of the x-, y- and z-component of the spatially averaged magnetization. Figure 6.2 shows that our results match those computed by the finite difference code *OOMMF* [DP99] available on the μMAG homepage [MUM]. Further, Figure 6.3 visualizes the magnetization at the time when the x-component of the spatially averaged magnetization first crosses zero.
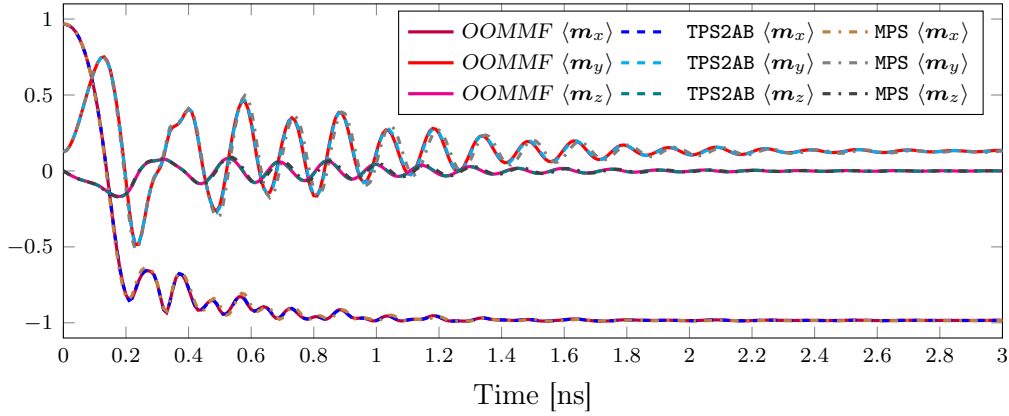
Figure 6.2: *μ*MAG standard problem #4 from Section 6.5.3: Time evolution of the spatially averaged magnetization components computed with *Commics* (`TPS2AB` and `MPS`) compared to the results of *OOMMF*.
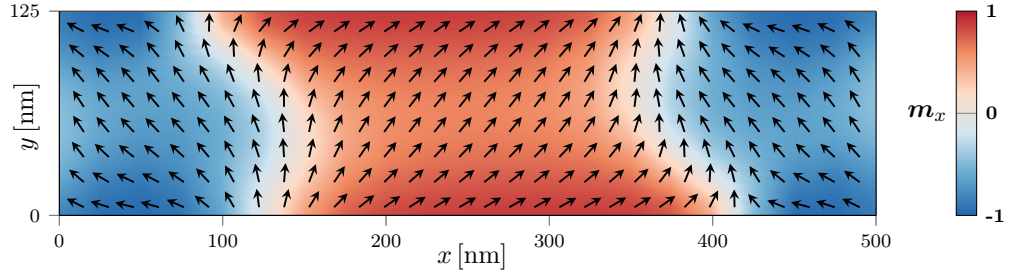

Figure 6.3: *μ*MAG standard problem #4 from Section 6.5.3: Snapshot of the magnetization when the *x*-component of the spatially averaged magnetization first crosses zero ($t = 138.2$ ps).

Meshing strategy and integrator

In Section 6.5.3, we considered a structured mesh and the `TPS2AB` integrator from [DPP$^+$20]. To compare the results, we additionally repeat the simulation on an unstructured mesh using the midpoint scheme: To simulate the dynamics on an unstructured mesh generated by *NGS*, we replace the definition of the geometry in Section 6.5.3 by

```
geo = Geometry(geometry=Cuboid(500, 125, 3), meshSize=3, scaling=1e-9)
```

This results in an unstructured mesh containing 48 792 elements, 16 682 vertices and 33 360 surface elements, which corresponds to an actual mesh size of 5.19 nm. To repeat the simulation using the midpoint scheme from [BP06], we replace the definition of the integrator in Section 6.5.3 and Section 6.5.3 by

```
sp4 = Integrator(scheme=MPS, geometry=geo, parameters=par)
```

Although the mesh size is close to the exchange length of the material (5.69 nm), qualitatively the results match those computed by *OOMMF* well; see Figure 6.2.

6.5.4 μMAG standard problem #5

The spintronic extensions of LLG from [ZL04, TNMS05] are the subject of the μMAG standard problem #5 [MUM]. The sample under consideration is a permalloy film with dimensions $100\,\text{nm} \times 100\,\text{nm} \times 10\,\text{nm}$ aligned with the x, y, and z axes of a Cartesian coordinate system, with origin at the center of the film. We consider the same material parameters as in Section 6.5.3 and $\alpha = 0.1$. The initial state is obtained by solving (6.3) with $\boldsymbol{T} \equiv \boldsymbol{0}$ and $\boldsymbol{m}^0(x,y,z) = (-y,x,R)/\sqrt{x^2 + y^2 + R^2}$ with $R = 10\,\text{nm}$ and maximal damping $\alpha = 1$ for $1\,\text{ns}$, which is a sufficiently long time for the system to reach equilibrium. Given $P\boldsymbol{J}_\text{e} = (1 \cdot 10^{12}, 0, 0)\,\text{A/m}^2$ and $\xi = 0.05$, we set $\boldsymbol{T}$ according to the expression in (6.7a). Then, we solve (6.3) with the relaxed magnetization configuration as initial condition for $8\,\text{ns}$, which turns out to be a sufficiently long time to reach the new equilibrium; see Figure 6.5 and Figure 6.4. For discretization, we use a constant time-step size $\Delta t = 0.1\,\text{ps}$ and consider a structured tetrahedral mesh of the given cuboid into cells of size $h_x \times h_y \times h_z$, which are decomposed into tetrahedra as depicted in Figure 6.1. The dimension of the cells is chosen as $h_x = h_y = h_z = 5/3\,\text{nm}$. This corresponds to $129\,600$ elements with diameter $h = 2.89\,\text{nm}$, $26\,047$ vertices, as well as $17\,280$ surface elements.

```python
from commics import *
# specify geometry and parameters
h = 5 / 3
geo = Geometry(geometry=Cuboid((-50,-50,-5), (50,50,5)), meshSize=(h,h,h), \
               structuredMesh=True, scaling=1.0e-09)
par = Parameters(A=1.3e-11, Ms=8.0e+05, K=0.0, gamma0=2.211e+05, alpha=1.0, \
                 spintronicsCoupling="zhangLi", g=2.0, P=1.0, \
                 Je=(1.0e+12, 0.0, 0.0), xi=0.05, \
                 timeStepSize=0.1e-12, T_start=-1.0e-09)
# space dependent initial condition (scaled domain); normalized automatically
from ngsolve import x, y
par.m0 = (-y, x, 10.0)
# define integrator and relax configuration
sp5 = Integrator(scheme=TPS2AB, geometry=geo, parameters=par)
sp5.Integrate(duration=1.0e-09, relax=True)
# set damping alpha and run simulation with specified current
sp5.SetParameter_alpha(0.1)
sp5.Integrate(duration=8.0e-09)
```
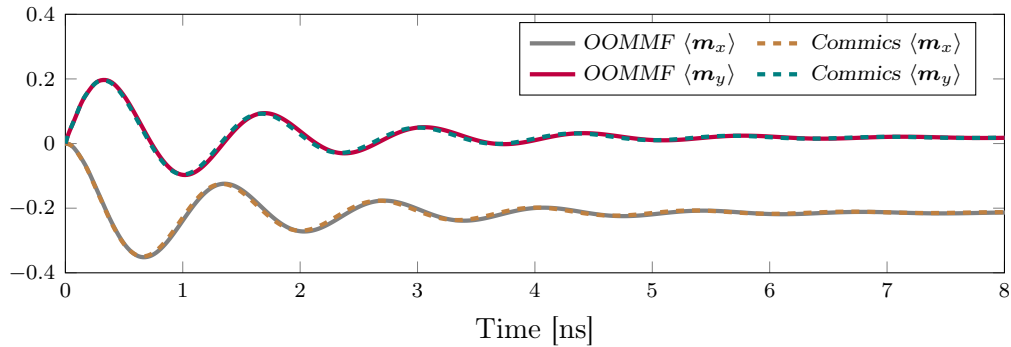

Figure 6.4: μMAG standard problem #5: Evolution of the spatially averaged x- and y-component of $\boldsymbol{m}$.
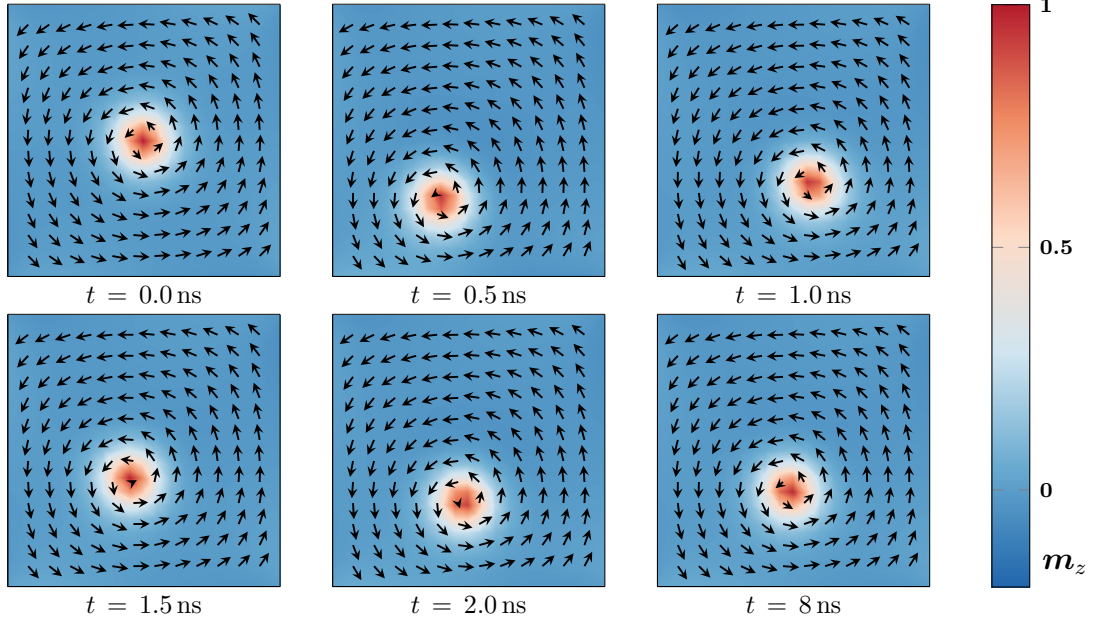
Figure 6.5: μMAG standard problem #5: Magnetization in the xy-plane viewed from the top at different times. Starting from the relaxed configuration at $t = 0.0$ ns, the vortex (red) follows a spiral-like motion. After $t = 8.0$ ns no further movements are observed.

6.5.5 Standard problem for ferromagnetic resonance simulations

Ferromagnetic resonance (FMR) is a well-established experimental technique for the study of ferromagnetic materials. A typical application of FMR consists in perturbing the magnetization of a system from its equilibrium by a sufficiently weak excitation and studying the induced magnetization dynamics, which is basically made of damped oscillations around the initial equilibrium. The resulting resonance frequencies and the eigenmodes of the system give some insights on the magnetic properties of the material and are used, e.g., for the experimental measurement of model parameters like the Gilbert damping constant or the saturation magnetization [Far98, MS05].

In this section, we compute with *Commics* a problem for FMR simulations recently proposed in [BBA$^+$17]. The computational domain is a cuboid of permalloy with dimensions 120 nm $\times$ 120 nm $\times$ 10 nm. The material parameters are the same as in Section 6.5.3. During the first stage, we set $\alpha = 1$ and consider a constant applied external field of magnitude $|\boldsymbol{H}_{\mathrm{ext}}| = 8 \cdot 10^4$ A/m pointing in the direction $(1, 0.715, 0)$. We initialize the system with a uniform ferromagnetic state $\boldsymbol{m}_h^0 \equiv (1, 0, 0)$ and let the system evolve for 5 ns. The resulting state is then used as initial condition for the second stage, in which we set $\alpha = 0.008$, change the direction of the applied external field to $(1, 0.7, 0)$ but keep $|\boldsymbol{H}_{\mathrm{ext}}| = 8 \cdot 10^4$ A/m , and let the system evolve to the new equilibrium for 20 ns. We consider a structured tetrahedral mesh of the given cuboid into cells of size $h_x \times h_y \times h_z$, which are decomposed into tetrahedra

as depicted in Figure 6.1. The dimension of the cells is chosen as $h_x = h_y = h_z = 2\,\text{nm}$. This corresponds to $108\,000$ elements, $22\,326$ vertices, as well as $16\,800$ surface elements, and yields a mesh size of $h = 3.46\,\text{nm}$. For time discretization, we use a constant time-step size $\Delta t = 0.1\,\text{ps}$. We compare our results obtained with *Commics* to those presented in [BBA$^+$17]. There, the authors use the finite difference code *OOMMF* [DP99] and investigate the evolution of the y-component of the spatially averaged magnetization, as well as its power spectrum S_y. The power spectrum is obtained by a discrete Fourier transform as described in [BBA$^+$17, Section 2.3.1]. Our results match well with those of [BBA$^+$17]; see Figure 6.6 and Figure 6.7.

```python
from commics import *
# specify geometry and parameters
geo = Geometry(geometry=Cuboid(120, 120, 10), meshSize=(2,2,2), \
               structuredMesh=True, scaling=1.0e-09)
par = Parameters(A=1.3e-11, Ms=8.0e+05, K=0.0, gamma0=2.210173e+05, \
                 alpha=1.0, timeStepSize=0.1e-12, T_start=-5.0e-9, \
                 m0=(0, 0, 1))
# define integrator, choose number of threads, specify folder for results
fmr = Integrator(scheme=TPS2AB, geometry=geo, parameters=par, numthreads=8)
fmr.SetResultsFolder("FMR_Result")
# obtain initial condition
amplitude = 80.0e+03
e = (1.0, 0.715, 0.0)
e_length = (sum(e[j]**2 for j in range(3)))**0.5
H_ext = (amplitude*e[0]/e_length, amplitude*e[1]/e_length, 0.0)
fmr.SetParameter_H_ext(H_ext)
fmr.Integrate(duration=5.0e-09)
# change direction of applied field and run simulation
fmr.SetParameter_alpha(0.008)
e = (1.0, 0.7, 0.0)
e_length = (sum(e[j]**2 for j in range(3)))**0.5
H_ext = (amplitude*e[0]/e_length, amplitude*e[1]/e_length, 0.0)
fmr.SetParameter_H_ext(H_ext)
fmr.Integrate(duration=20.0e-09)
```
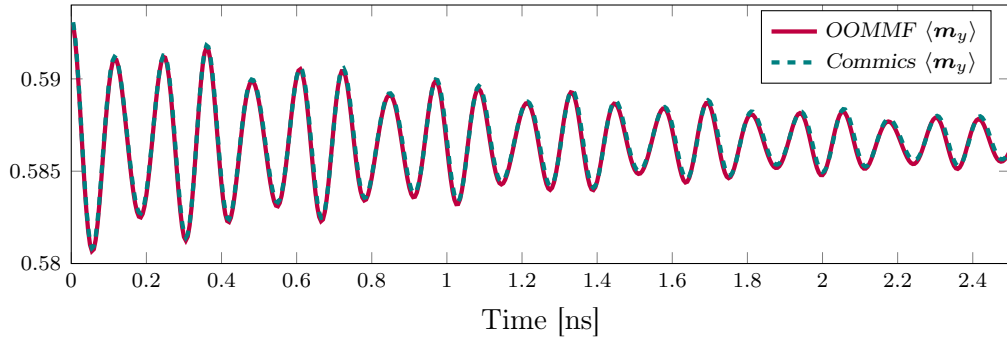

Figure 6.6: Ferromagnetic resonance simulation from Section 6.5.5: Time evolution of $\langle \boldsymbol{m}_y \rangle$ obtained with *Commics* compared to the results computed with *OOMMF*.
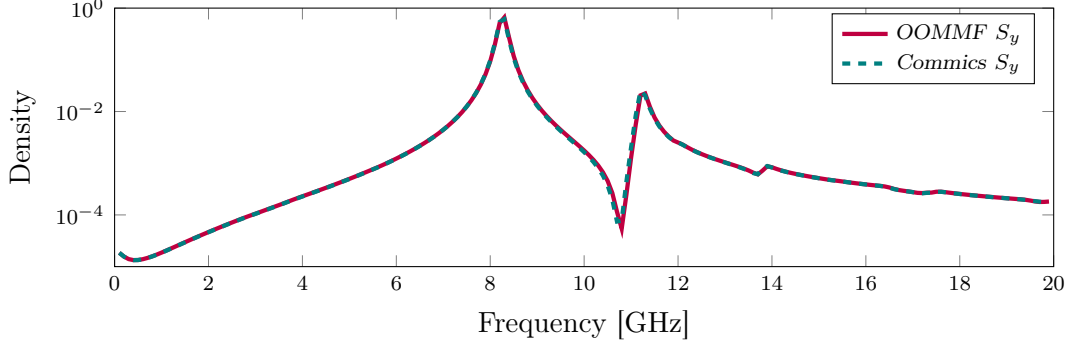
Figure 6.7: Ferromagnetic resonance simulation from Section 6.5.5: Power spectrum S_y obtained by discrete Fourier transform of the spatially averaged y-component of the magnetization $\langle \boldsymbol{m}_y \rangle$.

6.5.6 Current-induced dynamics of skyrmions in nanodisks

With this experiment, we aim to show how *Commics* can be used to numerically investigate the stability and the induced dynamics of magnetic skyrmions in helimagnetic materials in response to spin-polarized currents.

We consider a magnetic nanodisk of diameter 120 nm (x_1x_2-plane) and thickness $d =$ 10 nm (x_3-direction). We use the material parameters of iron-germanium (FeGe), i.e., $M_\mathrm{s} =$ $3.84 \cdot 10^5$ A/m, $A = 8.78 \cdot 10^{-12}$ J/m, $D = 1.58 \cdot 10^{-3}$ J/m^2, and $K = 0$ J/m^3; see, e.g., [BAB+17]. The initial condition for our experiment is obtained by relaxing a uniform out-of-plane ferromagnetic state $\boldsymbol{m}^0 \equiv (0,0,1)$ for 2 ns. For the relaxation process, we choose the large value $\alpha = 1$ for the Gilbert damping constant. The resulting relaxed state is the skyrmion depicted in Figure 6.9.
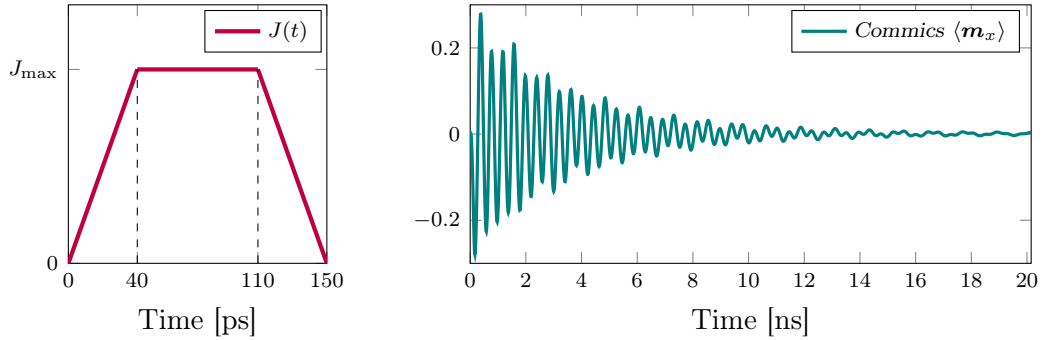
Figure 6.8: Simulation of the skyrmion dynamics from Section 6.5.6. Structure of the applied current pulse (left). Time evolution of the x-component of the spatially averaged magnetization (right).

Starting from this configuration, we apply a perpendicular spin-polarized current pulse $\boldsymbol{J}_\mathrm{e}(t) = (0,0,J(t))$ of maximum intensity $J_{\max} > 0$ for 150 ps; see Figure 6.8. To model the resulting spin transfer torque, we include $\boldsymbol{T}$ from (6.6). Then, we turn off the current

density and let the system evolve for 20 ns. In order to capture all possible excitation modes during the application of the pulse and the subsequent relaxation process, we set the value of the Gilbert damping constant to $\alpha = 0.002$, which is considerably smaller than the experimental value of $\alpha = 0.28$ measured for FeGe; see [BAB+17].

In Figure 6.8, we plot the time evolution of the first component of the spatially averaged magnetization of the sample after a current pulse with $J_{\max} = 1 \cdot 10^{12}\,\mathrm{A/m}^2$, $\boldsymbol{p} = (0, 1, 0)$, and $P = 0.4$. The induced dynamics is a damped precession of the skyrmion around the center of the sample; see Figure 6.9. We consider an unstructured tetrahedral mesh of the nanodisk generated by *NGS*. For a desired mesh size of 7.5 nm, the automatically generated mesh consists of 35 390 elements with maximum diameter $h = 6.1$ nm, 8436 vertices, as well as 9586 surface elements. We use a constant time-step size $\Delta t = 0.1$ ps.
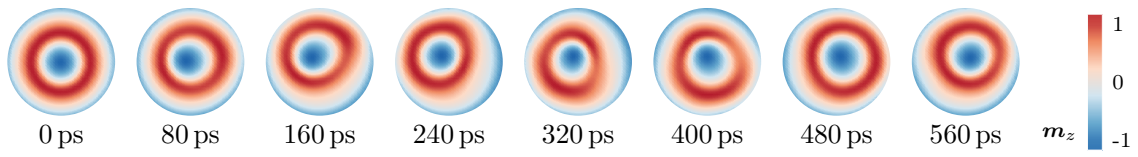

| 0 ps | 80 ps | 160 ps | 240 ps | 320 ps | 400 ps | 480 ps | 560 ps | $\boldsymbol{m}_z$ |

Figure 6.9: Snapshots of the skyrmion dynamics from Section 6.5.6: Magnetization in the xy-plane viewed from the top at different times. Starting from the relaxed configuration at $t = 0$ ps, the skyrmion is deflected from the center of the disk by a current pulse. Then, the skyrmion oscillates around the center of the disk with an observed period of approximately 400 ps. Due to damping, over the relaxation period of 20 ns the amplitude of the oscillations decreases to almost zero, and the initial equilibrium configuration is restored.

```python
from commics import *
# specify geometry and parameters
geo = Geometry(geometry=Disk(120.0, 10.0), meshSize=7.5, \
               forceNetgenMeshSize=True, scaling=1.0e-09)
par = Parameters(A=8.78e-12, Ms=3.84e+05, D=1.58e-03, dmCoupling="bulk", \
                 spintronicsCoupling="slonczewski", d=10.0e-09, P=0.4, \
                 p=(0.0, 1.0, 0.0), \
                 alpha=1.0, timeStepSize=1e-13, theta=1.0, m0=(0, 0, 1))
# define integrator, choose number of threads
helimag = Integrator(scheme=TPS1, geometry=geo, parameters=par, numthreads=8)
# choose to save magnetization as .vtk-file every X seconds, run simulation
helimag.RecordMagnetization(every=10.0e-12)
helimag.Integrate(duration=2.0e-9, relax=True)
# prepare JeMax, alpha, and points in time
helimag.SetParameter_alpha(0.002)
JeMax = 1.0e+12
T0, T1, T2, T3 = 0, 40.0e-12, 110.0e-12, 150.0e-12
# set increasing current, run simulation
helimag.SetParameter_Je(lambda t,x,y,z : (t-T0)/(T1-T0)*JeMax)
helimag.Integrate(duration=T1-T0)
# set constant current, run simulation
helimag.SetParameter_Je(JeMax)
helimag.Integrate(duration=T2-T1)
# set decreasing current, run simulation
helimag.SetParameter_Je(lambda t,x,y,z : (T3-t)/(T3-T2)*JeMax)
helimag.Integrate(duration=T3-T2)
# final relaxation
helimag.Integrate(duration=20.0e-9, relax=True)
```

6.5.7 Solver performance

In this section we provide details on the solver performance for the benchmark experiments from Sections 6.5.3–6.5.6. To solve the linear system for $\boldsymbol{v}_h^n$ arising in each time-step for tangent plane integrators, the efficient solution strategy from [KPP$^+$19], based on GMRES with a Jacobi-type preconditioner, is used. For the linear systems arising within the fixed-point iteration in each time-step of the midpoint scheme, diagonally preconditioned GMRES is employed. For both time integration schemes, the initial guess for GMRES is the solution vector of the system most recently solved by GMRES, i.e., the solution from the previous time-step for the tangent plane scheme, and the solution of the previous fixed-point iteration for the midpoint scheme. The performance for the different solvers and experiments is summarized in Table 6.1. All computations were performed on a desktop computer with 32 GB of RAM and an Intel Core i7-6700 CPU [Int] using 4 cores with a base frequency of 3.4 GHz.

Experiment	Integrator	Mesh			Iter. per time-step		Comp. time
		nodes	elements	surf. els.	fixed-point	GMRES	per time-step
μMAG #4	TPS2AB	58 170	228 528	78 936	–	74.1	4.5 s
	MPS	16 682	48 792	33 360	6.9	10.7	5.4 s
μMAG #5	TPS2AB	26 047	129 600	17 280	–	42.4	2.0 s
FMR	TPS2AB	22 326	108 000	16 800	–	63.4	1.5 s
Helimag	TPS1	8436	35 390	9586	–	57.2	0.41 s

Table 6.1: Experiments of Section 6.5.3 (μMAG #4), Section 6.5.4 (μMAG #5), Section 6.5.5 (FMR), and Section 6.5.6 (Helimag): Performance summary. The last three columns refer to the average cost of a time-step. Only time-steps after the initial relaxation phase are considered.

6.5.8 Coupling of *NGS* with *BEM++*

In this section we show how *NGS*, *BEM++*, and the *ngbem* module can be used to perform stray field computations following the approach proposed in [FK90]:

- *NGS* is used for mesh generation and the FEM problems.

- *BEM++* provides boundary integral operators and an interface for solving the BEM problem.

- *ngbem* provides the means to extract a boundary element mesh from the volume mesh together with a mapping between the corresponding degrees of freedom.

To test the procedure we consider the uniformly magnetized unit ball. Then, the stray field is given by $\boldsymbol{H}_{\mathrm{s}}(\boldsymbol{M}) \equiv -\boldsymbol{M}/3$ in Ω. The complete code is provided in the following Python script.

```
# Demonstration of the Fredkin-Koehler approach for stray field computations
# using ngsolve, ngbem, bempp to define mesh, femSpaces, GridFunctions
import ngsolve, ngbem, bempp.api, commics, numpy
geo = commics.Geometry(commics.UnitBall(), meshSize=0.3)
geo.GetReady()
mesh = geo.GetMesh()
Vh3 = ngsolve.VectorH1(mesh, order=1)
Vh = ngsolve.H1(mesh, order=1)
VhD = ngsolve.H1(mesh, order=1, dirichlet="bc_dirichlet")
m = ngsolve.GridFunction(Vh3)
m.components[0].vec.FV().NumPy()[:] = 1.0
u1 = ngsolve.GridFunction(Vh)
u2 = ngsolve.GridFunction(VhD)
hs = -u1.Deriv() - u2.Deriv()
# Neumann FEM problem
a1 = ngsolve.BilinearForm(Vh)
a1 += ngsolve.Laplace(1.0)
c1 = ngsolve.Preconditioner(a1, "local")
a1.Assemble()
solverU1 = ngsolve.CGSolver(mat=a1.mat, pre=c1.mat)
phi1 = Vh.TestFunction()
f1 = ngsolve.LinearForm(Vh)
f1 += ngsolve.SymbolicLFI(phi1.Deriv() * m)
measOmega = ngsolve.Integrate(ngsolve.CoefficientFunction(1.0), mesh)
# BEM problem
bemSpace, op_NgToBem = ngbem.H1_trace(Vh)
op_NgToBem.eliminate_zeros()
bemToNgIdx = numpy.full(op_NgToBem.shape[0], fill_value = Vh.ndof, dtype=int)
bemToNgIdx[op_NgToBem.row] = op_NgToBem.col
bem_K = bempp.api.operators.boundary.laplace.double_layer(bemSpace, bemSpace, \
                                                          bemSpace)
bem_I = bempp.api.operators.boundary.sparse.identity(bemSpace, bemSpace, \
                                                     bemSpace)
bemRhsOp = bem_K - 0.5*bem_I
bemLhs = bem_I
# Dirichlet FEM problem
a2 = ngsolve.BilinearForm(VhD)
a2 += ngsolve.Laplace(1.0)
c2 = ngsolve.Preconditioner(a2, "bddc")
a2.Assemble()
f2 = ngsolve.LinearForm(VhD)
f2.Assemble()
bvp2 = ngsolve.BVP(bf=a2, lf=f2, gf=u2, pre=c2)
# solve Neumann FEM problem
f1.Assemble()
u1.vec.data = solverU1 * f1.vec
u1.vec.FV().NumPy()[:] -= ngsolve.Integrate(u1, mesh) / measOmega
# solve BEM problem
bem_u1 = bempp.api.GridFunction(bemSpace, \
                                coefficients=u1.vec.FV().NumPy()[bemToNgIdx])
bem_rhs = bemRhsOp * bem_u1
bem_g = bempp.api.linalg.iterative_solvers.gmres(bemLhs, bem_rhs)[0]
# solve Dirichlet FEM problem
u2.vec.FV().NumPy()[bemToNgIdx] = bem_g.coefficients
bvp2.Do()
# check results
HS = ngsolve.GridFunction(Vh3)
for d in range(3): HS.components[d].Set(hs[d])
passed = all(abs(-1/3 - HS.components[0].vec.FV().NumPy()) < 0.01) \
         and all(abs(HS.components[1].vec.FV().NumPy()) < 0.01) \
         and all(abs(HS.components[2].vec.FV().NumPy()) < 0.01)
print("PASSED STRAYFIELD TEST:", passed)
```

Bibliography

[ABV14]   B. Ayuso de Dios, A. T. Barker, and P. S. Vassilevski. A combined preconditioning strategy for nonsymmetric systems. *SIAM J. Sci. Comput.*, 36(6):A2533–A2556, 2014.

[AD15]   F. Alouges and G. Di Fratta. Homogenization of composite ferromagnetic materials. *Proc. R. Soc. Lond. A*, 471(2182):20150365, 2015.

[ADMN21]   F. Alouges, A. De Bouard, B. Merlet, and L. Nicolas. Stochastic homogenization of the Landau–Lifshitz–Gilbert equation. *Stoch. Partial Differ. Equ. Anal. Comput.*, 9:789–818, 2021.

[AEB$^+$13]   C. Abert, L. Exl, F. Bruckner, A. Drews, and D. Suess. magnum.fe: A micromagnetic finite-element simulation code based on FEniCS. *J. Magn. Magn. Mater.*, 345:29–35, 2013.

[AES$^+$13]   C. Abert, L. Exl, G. Selke, A. Drews, and T. Schrefl. Numerical methods for the stray-field calculation: A comparison of recently developed algorithms. *J. Magn. Magn. Mater.*, 326:176–185, 2013.

[AFKL21]   G. Akrivis, M. Feischl, B. Kovács, and C. Lubich. Higher-order linearly implicit full discretization of the Landau-Lifshitz-Gilbert equation. *Math. Comp.*, 90(329):995–1038, 2021.

[AFM06]   E. Acerbi, I. Fonseca, and G. Mingione. Existence and regularity for mixtures of micromagnetic materials. *Proc. R. Soc. Lond. A*, 462(2072):2225–2243, 2006.

[AGL05]   J. Ahrens, B. Geveci, and C. Law. ParaView: An end-user tool for large-data visualization. In C. D. Hansen and C. R. Johnson, editors, *Visualization Handbook*, pages 717–731. Elsevier, 2005.

[Aha01]   A. Aharoni. *Introduction to the theory of ferromagnetism*. Oxford University Press, Oxford, second edition, 2001.

[AHP$^+$14]   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(6):639–654, 2014.

[AJ06]   F. Alouges and P. Jaisson. Convergence of a finite element discretization for the Landau–Lifshitz equation in micromagnetism. *Math. Models Methods Appl. Sci.*, 16(2):299–316, 2006.

[AKST14]   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(3):407–430, 2014.

[AKT12]   F. Alouges, E. Kritsikis, and J.-C. Toussaint. A convergent finite element approximation for Landau–Lifschitz–Gilbert equation. *Physica B*, 407(9):1345–1349, 2012.

[Alo08]   F. Alouges. A new finite element scheme for Landau–Lifchitz equations. *Discrete Contin. Dyn. Syst. Ser. S*, 1(2):187–196, 2008.

[An16]   R. An. Optimal error estimates of linearized Crank-Nicolson Galerkin method for Landau-Lifshitz equation. *J. Sci. Comput.*, 69(1):1–27, 2016.

[ARB+15]   C. Abert, M. Ruggeri, F. Bruckner, C. Vogler, G. Hrkac, D. Praetorius, and D. Suess. A three-dimensional spin-diffusion model for micromagnetics. *Scientific Reports*, 5:14855, 2015.

[AS92]   F. Alouges and A. Soyeur. On global weak solutions for Landau–Lifshitz equations: Existence and nonuniqueness. *Nonlinear Anal.*, 18(11):1071–1084, 1992.

[BAB+17]   M. Beg, M. Albert, M.-A. Bisotti, D. Cortés-Ortuño, W. Wang, R. Carey, M. Vousden, O. Hovorka, C. Ciccarelli, C. S. Spencer, C. H. Marrows, and H. Fangohr. Dynamics of skyrmionic states in confined helimagnetic nanostructures. *Phys. Rev. B*, 95(1):014433, 2017.

[Baň10]   L'. Baňas. An efficient multigrid preconditioner for Maxwell's equations in micromagnetism. *Math. Comput. Simulation*, 80(8):1657–1663, 2010.

[Bar05]   S. Bartels. Stability and convergence of finite-element approximation schemes for harmonic maps. *SIAM J. Numer. Anal.*, 43(1):220–238, 2005.

[Bar06]   S. Bartels. Constraint preserving, inexact solution of implicit discretizations of Landau–Lifshitz–Gilbert equations and consequences for convergence. *PAMM*, 6(1):19–22, 2006.

[Bar15]   S. Bartels. *Numerical methods for nonlinear partial differential equations.* Springer, Cham, 2015.

[Bar16]   S. Bartels. Projection-free approximation of geometrically constrained partial differential equations. *Math. Comp.*, 85(299):1033–1049, 2016.

[BBA+17]   A. Baker, M. Beg, G. Ashton, M. Albert, D. Chernyshenko, W. Wang, S. Zhang, M.-A. Bisotti, M. Franchin, C. L. Hu, R. Stamps, T. Hesjedal, and H. Fangohr. Proposal of a micromagnetic standard problem for ferromagnetic resonance simulations. *J. Magn. Magn. Mater.*, 421:428–439, 2017.

[BBF13]   D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications.* Springer, Berlin, 2013.

[BBNP14]   L'. Baňas, Z. Brzeźniak, M. Neklyudov, and A. Prohl. *Stochastic ferromagnetism: Analysis and numerics*, volume 58 of *Studies in Mathematics*. De Gruyter, 2014.

[BBP08]   L'. Baňas, S. Bartels, and A. Prohl. A convergent implicit finite element discretization of the Maxwell–Landau–Lifshitz–Gilbert equation. *SIAM J. Numer. Anal.*, 46(3):1399–1422, 2008.

[BBW+]   M.-A. Bisotti, M. Beg, W. Wang, M. Albert, D. Chernyshenko, D. Cortés-Ortuño, R. A. Pepper, M. Vousden, R. Carey, H. Fuchs, A. Johansen, G. Balaban, L. Leoni Breth, T. Kluyver, and H. Fangohr. FinMag: finite-element micromagnetic simulation tool (Version 0.1). Zenodo. `http://doi.org/10.5281/zenodo.1216011`. Accessed on February 4, 2022.

[BCOnP+18]   M.-A. Bisotti, D. Cortés-Ortuño, R. Pepper, W. Wang, M. Beg, T. Kluyver, and H. Fangohr. Fidimag – A finite difference atomistic and micromagnetic simulation package. *J. Open Res. Software*, 6(1), 2018.

[Ber96]   L. Berger. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B*, 54(13):9353–9358, 1996.

[BFF+14]   F. Bruckner, M. Feischl, T. Führer, P. Goldenits, M. Page, D. Praetorius, M. Ruggeri, and D. Suess. Multiscale modeling in micromagnetics: Existence of solutions and numerical integration. *Math. Models Methods Appl. Sci.*, 24(13):2627–2662, 2014.

[BH94]   A. Bogdanov and A. Hubert. Thermodynamically stable magnetic vortex states in magnetic crystals. *J. Magn. Magn. Mater.*, 138(3):255–269, 1994.

[BKP08]   S. Bartels, J. Ko, and A. Prohl. Numerical analysis of an explicit approximation scheme for the Landau–Lifshitz–Gilbert equation. *Math. Comp.*, 77(262):773–788, 2008.

[BL76]   J. Bergh and J. Löfström. *Interpolation space: An introduction*, volume 223 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1976.

[BOS+14]   S. Bance, H. Özelt, T. Schrefl, M. Winklhofer, G. Hrkac, G. Zimanyi, O. Gutfleisch, R.F.L. Evans, R.W. Chantrell, T. Shoji, M. Yano, N. Sakuma, A. Kato, and A. Manabe. High energy product in battenberg structured magnets. *Appl. Phys. Lett.*, 105(19):192401, 2014.

[BP06]   S. Bartels and A. Prohl. Convergence of an implicit finite element method for the Landau–Lifshitz–Gilbert equation. *SIAM J. Numer. Anal.*, 44(4):1405–1419, 2006.

[BPF17]   M. Beg, R. A. Pepper, and H. Fangohr. User interfaces for computational science: A domain specific language for OOMMF embedded in Python. *AIP Adv.*, 7(5):056025, 2017.

[BPP15]   L'. Baňas, M. Page, and D. Praetorius. A convergent linear finite element scheme for the Maxwell-Landau-Lifshitz-Gilbert equations. *Electron. Trans. Numer. Anal.*, 44:250–270, 2015.

[BPPR14]   L'. Baňas, M. Page, D. Praetorius, and J. Rochat. A decoupled and unconditionally convergent linear FEM integrator for the Landau-Lifshitz-Gilbert equation with magnetostriction. *IMA J. Numer. Anal.*, 34(4):1361–1385, 2014.

[BPS09]   L'. Baňas, A. Prohl, and M. Slodička. Modeling of thermally assisted magnetodynamics. *SIAM J. Numer. Anal.*, 47(1):551–574, 2009.

[BPS12]   L'. Baňas, A. Prohl, and M. Slodička. Numerical scheme for augmented Landau–Lifshitz equation in heat assisted recording. *J. Comput. Appl. Math.*, 236(18):4775–4787, 2012.

[BR01]   A. N. Bogdanov and U. K. Rößler. Chiral symmetry breaking in magnetic thin films and multilayers. *Phys. Rev. Lett.*, 87(3):037203, 2001.

[BRH93]   D. V. Berkov, K. Ramstöcck, and A. Hubert. Solving micromagnetic problems. Towards an optimal numerical method. *phys. stat. sol. (a)*, 137(1):207–225, 1993.

[Bro63]   W. F. Brown. *Micromagnetics.* Interscience Publishers, New York, 1963.

[BSH+08]   S. Bance, T. Schrefl, G. Hrkac, A. Goncharov, D. A. Allwood, and J. Dean. Micromagnetic calculation of spin wave propagation for magnetologic devices. *J. Appl. Phys.*, 103(7):07E735, 2008.

[BSP+17]   H. Brueckl, A. Satz, K. Pruegl, T. Wurft, S. Luber, W. Raberg, J. Zimmer, and D. Suess. Vortex magnetization state in a GMR spin-valve type field sensor. In *2017 IEEE International Magnetics Conference (INTERMAG)*, page 1, 2017.

[BY89]   A. Bogdanov and D. A. Yablonskii. Thermodynamically stable "vortices" in magnetically ordered crystals. The mixed state of magnets. *J. Exp. Theor. Phys.*, 68(1):101–103, 1989.

[CF01a]   G. Carbou and P. Fabrie. Regular solutions for Landau–Lifschitz equation in a bounded domain. *Differential Integral Equations*, 14(2):213–229, 2001.

[CF01b]   G. Carbou and P. Fabrie. Regular solutions for Landau-Lifschitz equation in $\mathbb{R}^3$. *Commun. Appl. Anal.*, 5(1):17–30, 2001.

[Cim08]   I. Cimrák. A survey on the numerics and computations for the Landau–Lifshitz equation of micromagnetism. *Arch. Comput. Methods Eng.*, 15(3):277–309, 2008.

[Cim09]   I. Cimrák. Convergence result for the constraint preserving mid-point scheme for micromagnetism. *J. Comput. Appl. Math.*, 228(1):238–246, 2009.

[CL98]   A. Crépieux and C. Lacroix. Dzyaloshinsky–Moriya interactions induced by symmetry breaking at a surface. *J. Magn. Magn. Mater.*, 182(3):341–349, 1998.

[CLL+11]   R. Chang, S. Li, M. V. Lubarda, B. Livshitz, and V. Lomakin. FastMag: Fast micromagnetic simulator for complex magnetic structures. *J. Appl. Phys.*, 109(7):07D358, 2011.

[CWX21]   J. Chen, C. Wang, and C. Xie. Convergence analysis of a second-order semi-implicit projection method for Landau-Lifshitz equation. *Appl. Numer. Math.*, 168:55–74, 2021.

[DD20]   E. Davoli and G. Di Fratta. Homogenization of chiral magnetic materials - A mathematical evidence of Dzyaloshinskii's predictions on helical structures. *J. Nonlinear Sci.*, 30:1229–1262, 2020.

[DIP20]   G. Di Fratta, M. Innerberger, and D. Praetorius. Weak-strong uniqueness for the Landau–Lifshitz–Gilbert equation in micromagnetics. *Nonlinear Anal. Real World Appl.*, 55:103122, 2020.

[DM17]   L. Döring and C. Melcher. Compactness results for static and dynamic chiral skyrmions near the conformal limit. *Calc. Var. Partial Differential Equations*, 56:60, 2017.

[DMRS20]   G. Di Fratta, C. B. Muratov, F. N. Rybakov, and V. V Slastikov. Variational principles of micromagnetics revisited. *SIAM J. Math. Anal.*, 52(4):3580–3599, 2020.

[DP99]   M. J. Donahue and D. G. Porter. OOMMF user's guide, Version 1.0. Interagency Report NISTIR 6376, National Institute of Standards and Technology, Gaithersburg, MD, 1999.

[DPP+20]   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(4):2802–2838, 2020.

[DS14]   E. Dumas and F. Sueur. On the weak solutions to the Maxwell–Landau–Lifshitz equations and to the Hall-Magneto-Hydrodynamic equations. *Comm. Math. Phys.*, 330(3):1179–1225, 2014.

[DSM05]   M. D'Aquino, C. Serpico, and G. Miano. Geometrical integration of Landau–Lifshitz–Gilbert equation based on the mid-point rule. *J. Comput. Phys.*, 209(2):730–753, 2005.

[Dzy58]   I. Dzyaloshinskii. A thermodynamic theory of 'weak' ferromagnetism of antiferromagnetics. *J. Phys. Chem. Solids*, 4(4):241–255, 1958.

[Eva10]   L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, second edition, 2010.

[Far98]   M. Farle. Ferromagnetic resonance of ultrathin metallic layers. *Rep. Prog. Phys.*, 61(7):755, 1998.

[FBT+16] G. Finocchio, F. Büttner, R. Tomasello, M. Carpentieri, and M. Kläui. Magnetic skyrmions: from fundamental to applications. *J. Phys. D: Appl. Phys.*, 49(42):423001, 2016.

[FCH+16] S. Fu, W. Cui, M. Hu, R. Chang, M. J. Donahue, and V. Lomakin. Finite-difference micromagnetic solvers with the object-oriented micromagnetic framework on graphics processing units. *IEEE Trans. Magn.*, 52(4):1–9, 2016.

[FCS13]   A. Fert, V. Cros, and J. Sampaio. Skyrmions on the track. *Nat. Nanotechnol.*, 8(3):152–156, 2013.

[fem]   FEMME. https://www.suessco.com/simulations/. Accessed on February 4, 2022.

[FFBF07] T. Fischbacher, M. Franchin, G. Bordignon, and H. Fangohr. A systematic approach to multiphysics extensions of finite-element-based micromagnetic simulations: Nmag. *IEEE Trans. Magn.*, 43(6):2896–2898, 2007.

[FK90]   D. R. Fredkin and T. R. Koehler. Hybrid method for computing demagnetization fields. *IEEE Trans. Magn.*, 26(2):415–417, 1990.

[FKO+17] J. Fischbacher, A. Kovacs, H. Özelt, M. Gusenbauer, T. Schrefl, L. Exl, D. Givord, N. M. Dempsey, G. Zimanyi, M. Winklhofer, et al. On the limits of coercivity in permanent magnets. *Appl. Phys. Lett.*, 111(7):072404, 2017.

[FRC17]   A. Fert, N. Reyren, and V. Cros. Magnetic skyrmions: advances in physics and potential applications. *Nat. Rev. Mater.*, 2:17031, 2017.

[FT17a]   M. Feischl and T. Tran. The Eddy Current-LLG equations: FEM-BEM coupling and a priori error estimates. *SIAM J. Numer. Anal.*, 55(4):1786–1819, 2017.

[FT17b]   M. Feischl and T. Tran. Existence of regular solutions of the Landau-Lifshitz-Gilbert equation in 3D with natural boundary conditions. *SIAM J. Math. Anal.*, 49(6):4470–4490, 2017.

[Gao14]   H. Gao. Optimal error estimates of a linearized backward Euler FEM for the Landau-Lifshitz equation. *SIAM J. Numer. Anal.*, 52(5):2574–2593, 2014.

[GC07]   C. J. García-Cervera. Numerical micromagnetics: A review. *Bol. Soc. Esp. Mat. Apl. SeMA*, 39:103–135, 2007.

[GH93]   B. Guo and M.-C. Hong. The Landau–Lifshitz equation of the ferromagnetic spin chain and harmonic maps. *Calc. Var. Partial Differential Equations*, 1(3):311–334, 1993.

[Gil55]   T. L. Gilbert. A Lagrangian formulation of the gyromagnetic equation of the magnetization fields. *Phys. Rev.*, 100:1243, 1955. Abstract only.

[Gil04]   T. L. Gilbert. A phenomenological theory of damping in ferromagnetic materials. *IEEE Trans. Magn.*, 40(6):3443–3449, 2004.

[GO95]   M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70(2):163–180, 1995.

[Gol12]   P. Goldenits. *Konvergente numerische Integration der Landau–Lifshitz–Gilbert Gleichung*. PhD thesis, TU Wien, 2012. In German.

[GR86]   V. Girault and P.-A. Raviart. *Finite element methods for Navier–Stokes equations: Theory and algorithms*, volume 5 of *Springer Series in Computational Mathematics*. Springer, 1986.

[Hac99]   W. Hackbusch. A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices. *Computing*, 62(2):89–108, 1999.

[Her01]   R. Hertel. Micromagnetic simulations of magnetostatically coupled nickel nanowires. *J. Appl. Phys.*, 90(11):5752–5758, 2001.

[HK14]   R. Hertel and A. Kákay. Hybrid finite-element/boundary-element method to calculate oersted fields. *J. Magn. Magn. Mater.*, 369:189–196, 2014.

[HKBB15]   G. Hrkac, P. S. Keatley, M. T. Bryan, and K. Butler. Magnetic vortex oscillators. *J. Phys. D: Appl. Phys.*, 48(45):453001, 2015.

[HKK$^+$16] D.-S. Han, N.-H. Kim, J.-S. Kim, Y. Yin, J.-W. Koo, J. Cho, S. Lee, M. Kläui, H. J. M. Swagten, B. Koopmans, and C.-Y. You. Asymmetric hysteresis for probing Dzyaloshinskii–Moriya interaction. *Nano Lett.*, 16(7):4438–4446, 2016.

[HPP$^+$19] 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(3):1329–1368, 2019.

[HS98]   A. Hubert and R. Schäfer. *Magnetic domains: The analysis of magnetic microstructures*. Springer, 1998.

[HvBM$^+$11] S. Heinze, K. von Bergmann, M. Menzel, J. Brede, A. Kubetzka, R. Wiesendanger, G. Bihlmayer, and S. Blugel. Spontaneous atomic-scale magnetic skyrmion lattice in two dimensions. *Nat. Phys.*, 7(9):713–718, 2011.

[Int]   Intel. Intel$^\circledR$ Core i7-6700 Processor. `https://ark.intel.com/content/www/us/en/ark/products/88196/intel-core-i7-6700-processor-8m-cache-up-to-4-00-ghz.html`. February 3, 2022.

[JOP⁺]   E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python. `http://www.scipy.org/`, 2001–. Accessed on February 4, 2022.

[KL17]   E. Kim and K. Lipnikov. The mimetic finite difference method for the Landau-Lifshitz equation. *J. Comput. Phys.*, 328:109–130, 2017.

[KOSS16] A. Kovacs, H. Özelt, M.E. Schabes, and T. Schrefl. Numerical optimization of writer and media for bit patterned magnetic recording. *J. Appl. Phys.*, 120(1):013902, 2016.

[KP06]   M. Kruzik and A. Prohl. Recent developments in the modeling, analysis, and numerics of ferromagnetism. *SIAM Rev.*, 48(3):439–483, 2006.

[KP15]   S. Komineas and N. Papanicolaou. Skyrmion dynamics in chiral ferromagnets. *Phys. Rev. B*, 92(6):064412, 2015.

[KPP⁺18] J. Kraus, C.-M. Pfeiler, D. Praetorius, M. Ruggeri, and B. Stiftner. Iterative solution and preconditioning for the tangent plane scheme in computational micromagnetics. arXiv:1808.10281, 2018.

[KPP⁺19] J. Kraus, C.-M. Pfeiler, D. Praetorius, M. Ruggeri, and B. Stiftner. Iterative solution and preconditioning for the tangent plane scheme in computational micromagnetics. *J. Comput. Phys.*, 398:108866, 2019.

[Kro07]   H. Kronmüller. General micromagnetic theory. In H. Kronmüller, P. Parkin, J. E. Miltat, and M. R. Scheinfein, editors, *Handbook of Magnetism and Advanced Magnetic Materials*. John Wiley & Sons, Ltd, 2007.

[KSH⁺16] P. S. Keatley, S. R. Sani, G. Hrkac, S. M. Mohseni, P. Dürrenfeld, T. H. J. Loughran, J. Åkerman, and R. J. Hicken. Direct observation of magnetization dynamics generated by nanocontact spin-torque vortex oscillators. *Phys. Rev. B*, 94:060402(R), 2016.

[KSH⁺17] P. S. Keatley, S. R. Sani, G. Hrkac, S. M. Mohseni, P. Dürrenfeld, J. Åkerman, and R. J. Hicken. Imaging magnetisation dynamics in nano-contact spin-torque vortex oscillators exhibiting gyrotropic mode splitting. *J. Phys. D: Appl. Phys.*, 50(16):164003, 2017.

[KW18]   E. Kim and J. Wilkening. Convergence of a mass-lumped finite element method for the Landau–Lifshitz equation. *Quart. Appl. Math.*, 76:383–405, 2018.

[KWH10]   A. Kakay, E. Westphal, and R. Hertel. Speedup of FEM micromagnetic simulations with graphical processing units. *IEEE Trans. Magn.*, 46(6):2303–2306, 2010.

[LL35]   L. Landau and E. Lifshitz. On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. *Phys. Zeitsch. der Sow.*, 8:153–168, 1935.

[LL08]   L. Landau and E. Lifschitz. On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. *Ukr. J. Phys.*, 53, 2008. Special Issue, Reprinted from Phys. Zeitsch. der Sow. 8, pp. 153–169 (1935).

[LPPT15]   K.-N. Le, M. Page, D. Praetorius, and T. Tran. On a decoupled linear FEM integrator for eddy-current-LLG. *Appl. Anal.*, 94(5):1051–1067, 2015.

[LT13]   K.-N. Le and T. Tran. A convergent finite element approximation for the quasi-static Maxwell-Landau-Lifshitz-Gilbert equations. *Comput. Math. Appl.*, 66(8):1389–1402, 2013.

[LW08]   F. Lin and C. Wang. *The analysis of harmonic maps and their heat flows.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.

[mag]   magnum.fd. `http://micromagnetics.org/magnum.fd`. Accessed on February 4, 2022.

[MBJ⁺09]   S. Mühlbauer, B. Binz, F. Jonietz, C. Pfleiderer, A. Rosch, A. Neubauer, R. Georgii, and P. Böni. Skyrmion lattice in a chiral magnet. *Science*, 323(5916):915–919, 2009.

[MD07]   J. E. Miltat and M. J. Donahue. Numerical micromagnetics: Finite difference methods. In H. Kronmüller, P. Parkin, J. E. Miltat, and M. R. Scheinfein, editors, *Handbook of Magnetism and Advanced Magnetic Materials*. John Wiley & Sons, Ltd, 2007.

[Mel05]   C. Melcher. Existence of partially regular solutions for Landau–Lifshitz equations in $\mathbb{R}^3$. *Comm. Partial Differential Equations*, 30(4):567–587, 2005.

[Mel14]   C. Melcher. Chiral skyrmions in the plane. *Proc. R. Soc. A*, 470(2172), 2014.

[mic]   MicroMagnum. `https://github.com/MicroMagnum/MicroMagnum`. Accessed on February 4, 2022.

[Mor60]   T. Moriya. Anisotropic superexchange interaction and weak ferromagnetism. *Phys. Rev.*, 120(91):91, 1960.

[MPPR21]   N. Mauser, C.-M. Pfeiler, D. Praetorius, and M. Ruggeri. Unconditional well-posedness and IMEX improvement of a family of predictor-corrector methods in micromagnetics. preprint, arXiv:2112.00451, 2021.

[MS05]   R. D. McMichael and M. S. Stiles. Magnetic normal modes of nanoelements. *J. Appl. Phys.*, 97(10):10J901, 2005.

[MS17]   C. B. Muratov and V. V. Slastikov. Domain structure of ultrathin ferromagnetic elements in the presence of Dzyaloshinskii–Moriya interaction. *Proc. R. Soc. A*, 473(2197), 2017.

[MSOS12]   A. Makarov, V. Sverdlov, D. Osintsev, and S. Selberherr. Fast switching in magnetic tunnel junctions with two pinned layers: Micromagnetic modeling. *IEEE Trans. Magn.*, 48(4):1289–1292, 2012.

[MUM]   NIST micromagnetic modeling activity group (μMAG) website. `http://www.ctcms.nist.gov/~rdm/mumag.html`. Accessed on February 3, 2022.

[Nep91]   S. V. Nepomnyaschikh. Mesh theorems on traces, normalizations of function traces and their inversion. *Soviet J. Numer. Anal. Math. Modelling*, 6(3):223–242, 1991.

[ngs]   Netgen/NGSolve Finite Element Library. `https://ngsolve.org/`. Accessed on February 3, 2022.

[NS96]   S. G. Nash and A. Sofer. Preconditioning reduced matrices. *SIAM J. Matrix Anal. Appl.*, 17(1):47–68, 1996.

[NT13]   N. Nagaosa and Y. Tokura. Topological properties and dynamics of magnetic skyrmions. *Nat. Nanotechnol.*, 8(12):899–911, 2013.

[OOM]   The Object Oriented MicroMagnetic Framework (OOMMF) project at ITL/NIST. `http://math.nist.gov/oommf/`. Accessed on February 3, 2022.

[OS18]   L. N. Olson and J. B. Schroder. PyAMG: Algebraic multigrid solvers in Python v4.0. `https://github.com/pyamg/pyamg`, 2018. Accessed on February 4, 2022.

[Pfe]   C.-M. Pfeiler. Commics – A Python module for Computational Micromagnetics. `https://gitlab.asc.tuwien.ac.at/cpfeiler/commics`. Accessed on February 4, 2022.

[PHT08]   S. S. P. Parkin, M. Hayashi, and L. Thomas. Magnetic domain-wall racetrack memory. *Science*, 320(5873):190–194, 2008.

[Pra04]   D. Praetorius. Analysis of the operator Δ^{-1}div arising in magnetic models. *Z. Anal. Anwend.*, 23(3):589–605, 2004.

[Pro01]   A. Prohl. *Computational micromagnetism*. Teubner, Wiesbaden, 2001.

[PRS⁺18a] C.-M. Pfeiler, M. Ruggeri, B. Stiftner, L. Exl, M. Hochsteger, G. Hrkac, J. Schöberl, N. Mauser, and D. Praetorius. Computational micromagnetics with Commics. preprint, arXiv:1812.05931, 2018.

[PRS18b]   D. Praetorius, M. Ruggeri, and B. Stiftner. Convergence of an implicit-explicit midpoint scheme for computational micromagnetics. *Comput. Math. Appl.*, 75(5):1719–1738, 2018.

[PRS⁺20]   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.*, 248:106965, 2020.

[QSS07]   A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*, volume 37 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, second edition, 2007.

[RBP06]   U. K. Rößler, A. N. Bogdanov, and C. Pfleiderer. Spontaneous skyrmion ground states in magnetic metals. *Nature*, 442(7104):797–801, 2006.

[RHM⁺13]   N. Romming, C. Hanneken, M. Menzel, J. E. Bickel, B. Wolter, K. von Bergmann, A. Kubetzka, and R. Wiesendanger. Writing and deleting single magnetic skyrmions. *Science*, 341(6146):636–639, 2013.

[Rie]   A. Rieder. Coupling NGsolve to BEM++. `https://github.com/arieder/ngbem`. Accessed on February 4, 2022.

[Rug16]   M. Ruggeri. *Coupling and numerical integration of the Landau–Lifshitz–Gilbert equation*. PhD thesis, TU Wien, Institute for Analysis and Scientific Computing, 2016.

[Saa03]   Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.

[SAON⁺13]   H. Sepehri-Amin, T. Ohkubo, S. Nagashima, M. Yano, T. Shoji, A. Kato, T. Schrefl, and K. Hono. High-coercivity ultrafine-grained anisotropic Nd–Fe–B magnets processed by hot deformation and the Nd–Cu grain boundary diffusion process. *Acta Mater.*, 61(17):6622–6634, 2013.

[ŚBA⁺15]   W. Śmigaj, T. Betcke, S. Arridge, J. Phillips, and M. Schweiger. Solving boundary integral problems with BEM++. *ACM Trans. Math. Software*, 41(2):Art. 6, 40, 2015.

[Sch97]   J. Schöberl. NETGEN An advancing front 2D/3D-mesh generator based on abstract rules. *Comput. Vis. Sci.*, 1(1):41–52, 1997.

[SCR⁺13]   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(11):839–844, 2013.

[SFS⁺03]   W. Scholz, J. Fidler, T. Schrefl, D. Suess, R. Dittrich, H. Forster, and V. Tsiantos. Scalable parallel micromagnetic solvers for magnetic nanostructures. *Comput. Mater. Sci.*, 28(2):366–383, 2003.

[SHB⁺07]   T. Schrefl, G. Hrkac, S. Bance, D. Suess, O. Ertl, and J. Fidler. Numerical methods in micromagnetics (finite element method). In H. Kronmüller, P. Parkin, J. E. Miltat, and M. R. Scheinfein, editors, *Handbook of Magnetism and Advanced Magnetic Materials*. John Wiley & Sons, Ltd, 2007.

[Slo96]   J. C. Slonczewski. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mat.*, 159(1–2):L1–L7, 1996.

[SS86]   Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, 1986.

[SSB86]   P. L. Sulem, C. Sulem, and C. Bardos. On the continuous limit for a system of classical spins. *Comm. Math. Phys.*, 107(3):431–454, 1986.

[Sta97]   G. Starke. Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems. *Numer. Math.*, 78(1):103–117, 1997.

[STS+02]   D. Suess, V. Tsiantos, T. Schrefl, J. Fidler, W. Scholz, H. Forster, R. Dittrich, and J. J. Miles. Time resolved micromagnetics using a preconditioned time integration method. *J. Magn. Magn. Mater.*, 248(2):298–311, 2002.

[SVA+15]   D. Suess, C. Vogler, C. Abert, F. Bruckner, R. Windl, L. Breth, and J. Fidler. Fundamental limits in heat-assisted magnetic recording and methods to overcome it with exchange spring structures. *J. Appl. Phys.*, 117(16):163913, 2015.

[SZZS00]   R. Skomski, H. Zeng, M. Zheng, and D. J. Sellmyer. Magnetic localization in transition-metal nanowires. *Phys. Rev. B*, 62(6):3900, 2000.

[Tho06]   V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer, second edition, 2006.

[TMZ+14]   R. Tomasello, E. Martinez, R. Zivieri, L. Torres, M. Carpentieri, and G. Finocchio. A strategy for the design of skyrmion racetrack memories. *Sci. Rep.*, 4:6784, 2014.

[TNMS05]   A. Thiaville, Y. Nakatani, J. Miltat, and Y. Suzuki. Micromagnetic understanding of current-driven domain wall motion in patterned nanowires. *EPL (Europhysics Letters)*, 69(6):990, 2005.

[TRJF12]   A. Thiaville, S. Rohart, V. Jué, É. Cros, and A. Fert. Dynamics of Dzyaloshinskii domain walls in ultrathin magnetic films. *Europhys. Lett.*, 100(5):57002, 2012.

[Vis85]   A. Visintin. On Landau–Lifshitz' equations for ferromagnetism. *Japan J. Appl. Math.*, 2(1):69–84, 1985.

[VLD+14]   A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. Van Waeyenberge. The design and verification of MuMax3. *AIP Adv.*, 4(10):107133, 2014.

[WGCE01]   X.-P. Wang, C. J. García-Cervera, and W. E. A Gauss–Seidel projection method for micromagnetics simulations. *J. Comput. Phys.*, 171(1):357–372, 2001.

[Wie16]   R. Wiesendanger. Nanoscale magnetic skyrmions in metallic films and multilayers: a new twist for spintronics. *Nat. Rev. Mater.*, 1:16044, 2016.

[XGCW⁺20] C. Xie, C. J. García-Cervera, C. Wang, Z. Zhou, and J. Chen. Second-order semi-implicit projection methods for micromagnetics simulations. *J. Comput. Phys.*, 404:109104, 2020.

[ZL04] S. Zhang and Z. Li. Roles of nonequilibrium conduction electrons on the magnetization dynamics of ferromagnets. *Phys. Rev. Lett.*, 93(12):127204, 2004.

Curriculum Vitae

Personal information

Name	**Carl-Martin Pfeiler**
Geburtsdatum	November 11, 1991
Geburtsort	Innsbruck, Austria
Citizenship	Austrian
Email	`carl-martin.pfeiler@asc.tuwien.ac.at`
Webpage	`https://www.asc.tuwien.ac.at/~cpfeiler`

Education

03/2022 (expected)	**TU Wien, Vienna, Austria** Doktor der Technischen Wissenschaften (PhD) Supervisor: Prof. Dirk Praetorius
01/2018	**TU Wien, Vienna, Austria** Diplom Ingenieur (MSc) in Technical Mathematics Supervisor: Prof. Dirk Praetorius
11/2015	**TU Wien, Vienna, Austria** Bachelor of Science in Technical Mathematics Supervisor: Prof. Dirk Praetorius
06/2010	**BG/BRG Sillgasse, Innsbruck, Austria** Matura

Prizes

2021	**Best Paper Award 2020** (1.000 €) of the Faculty of Mathematics and Geoinformation at TU Wien for the work *Dörfler marking with minimal cardinality is a linear complexity problem.*

Publications

2022   G. Di Fratta, C.-M. Pfeiler, D. Praetorius and M. Ruggeri. *The mass-lumped midpoint scheme for computational micromagnetics: Newton linearization and application to magnetic skyrmion dynamics*, in preperation.

2021   N. Mauser, C.-M. Pfeiler, D. Praetorius and M. Ruggeri. *Unconditional well-posedness and IMEX improvement of a family of predictor-corrector methods in micromagnetics*, in submission, preprint available at arXiv: 2112.00451.

2020   C.-M. Pfeiler and D. Praetorius. *Dörfler marking with minimal cardinality is a linear complexity problem*, Mathematics of Computation 89: 2735–2752. doi: 10.1090/mcom/3553.

2020   G. Di Fratta, C.-M. Pfeiler, D. Praetorius, M. Ruggeri and B. Stiftner. *Linear second-order implicit-explicit in time integration of the (Eddy-currents-) Landau–Lifshitz–Gilbert equation.* IMA Journal of Numerical Analysis, 40: 2802–2838. doi: 10.1093/imanum/drz046.

2020   C.-M. Pfeiler, M. Ruggeri, B. Stiftner, L. Exl, M. Hochsteger, G. Hrkac, J. Schöberl, N. Mauser and D. Praetorius. *Computational micromagnetics with Commics.* Computer Physics Communications, 248: 106965. doi: 10.1016/j.cpc.2019.106965.

2019   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.* Advances in Computational Mathematics, 45: 1329–1368. doi: 10.1007/s10444-019-09667-z.

2019   J. Kraus, C.-M. Pfeiler, D. Praetorius, M. Ruggeri and B. Stiftner. *Iterative solution and preconditioning for the tangent plane scheme in computational micromagnetics.* Journal of Computational Physics, 398: 108866. doi: 10.1016/j.jcp.2019.108866.

2016   P. Lederer, C.-M. Pfeiler, C. Wintersteiger and C. Lehrenfeld. *Higher order unfitted FEM for Stokes interface problems.* Proceedings in Applied Mathematics and Mechanics, 16: 7–10. doi: 10.1002/pamm.201610003.

Conference contributions

DMV 2019 Karlsruhe, Germany	Talk at *DMV-Jahrestagung* Dörfler marking with minimal cardinality is a linear complexity problem
RMMM 2019 Wien, Austria	Talk at *Reliable Methods of Mathematical Modeling* Dörfler marking with minimal cardinality is a linear complexity problem
HMM 2019 Heraklion, Crete	Poster at *12th int. Symposium on Hysteresis Modeling and Micromagnetics* Computational studies of nonlinear skyrmion dynamics
ANADay 2019 Graz, Austria	Talk at *Austrian Numerical Analysis Day* Dörfler marking with minimal cardinality is a linear complexity problem
WONAPDE 2019 Concepción, Chile	Talk at *Sixth Chilean Workshop on Numerical Analysis of Partial Differential Equations* Dörfler marking with minimal cardinality is a linear complexity problem
MANA 2018 Wien, Austria	Poster at *Micromagnetics: Analysis, Numerics, Applications* Computational micromagnetics with Commics
NGS 2018 Göttingen	Talk at *2nd NGSolve User Meeting* Computational micromagnetics in NGSolve

---

Wien, am 10. Februar 2022

_____
Carl-Martin Pfeiler