Karlsruhe Institute of Technology

Space and Time Adaptivity for the Landau–Lifshitz–Gilbert Equation

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

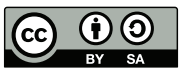
von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Stefan Karch

Tag der mündlichen Prüfung:	22. April 2026
Referent:	Prof. Dr. Willy Dörfler
Korreferent:	Prof. Dr. Christian Wieners



Acknowledgments

Zuallererst möchte ich mich bei meinem Betreuer Prof. Dr. Willy Dörfler bedanken, der mich in den vergangenen Jahren stets unterstützt hat. Wenn ich mit einer Frage zu ihm kam, erhielt ich nicht nur eine passende Antwort, sondern auch mindestens zwei neue Ideen mit auf den Weg. Neben der fachlichen Betreuung danke ich ihm zudem für die hervorragende Zusammenarbeit in der Lehre, durch die ich auch persönlich viel gelernt habe.

Des Weiteren gilt mein Dank Prof. Dr. Christian Wieners für die Übernahme des Zweitgutachtens sowie für seine wertvollen Kommentare, die sehr positiv zur Qualität dieser Arbeit beigetragen haben.

Darüber hinaus möchte ich mich auch bei Prof. Dr. Michael Feischl für den Forschungsaufenthalt in Wien sowie für den bereichernden fachlichen Austausch bedanken.

Ein großer Dank geht auch an Isabell, Joshua und Tim für das hilfreiche Korrekturlesen von Teilen dieser Dissertation. Ebenso danke ich Raphael für die wertvolle Unterstützung gerade zu Beginn meiner Dissertation sowie Jan dafür, dass er sich Zeit genommen hat, mir einen guten Einstieg in das Thema zu ermöglichen. Zusätzlich möchte ich mich bei Dustin, Johanna, Mariia, Mehdi, Michael und Tino aus meiner Arbeitsgruppe bedanken. Danke für die schöne Zeit und die vielen gemeinsamen Mittagessen.

Außerdem danke ich Nathalie für die hervorragende organisatorische Unterstützung und ihre große Hilfsbereitschaft. Mein Dank gilt zudem Mathias für die konstruktive Zusammenarbeit in der Lehre sowie allen Tutorinnen und Tutoren, die uns dabei unterstützt haben.

Ein großer Dank geht auch an Daniel, Lukas, Selina und Tim, die mich nicht nur im SFB, sondern bereits während meines gesamten Studiums begleitet haben. Ebenso danke ich allen Kolleginnen und Kollegen aus dem SFB für die angenehme Zeit bei den Jahrestreffen sowie im gesamten Arbeitsalltag.

Darüber hinaus danke ich auch allen, die ich hier nicht explizit aufgeführt habe. Zuletzt möchte ich mich bei meiner Familie bedanken, auf deren Unterstützung ich mich stets verlassen konnte.

I gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173.

Abstract

The Landau–Lifshitz–Gilbert (LLG) equation serves as the fundamental model to describe magnetization dynamics in ferromagnetic materials. Due to its highly nonlinear structure, which includes a non-convex constraint preserving the magnetization length at every point in space and time, the equation poses significant challenges for the numerical analysis. In particular, small damping constants complicate the mathematical stability analysis and the efficient numerical simulation considerably. Since highly localized phenomena, such as domain walls in space or rapid switching events in time, occur in practice, the use of adaptive methods is well-suited to resolve these local structures.

In this work, we investigate a tangent plane scheme for solving the LLG equation that enforces the intrinsic normalization of the magnetization in a weak sense. We employ finite elements for spatial discretization and utilize linearly implicit backward differentiation formulas (BDF) for time integration. While the higher-order convergence of this method is already established, the focus of this thesis lies on the investigation and derivation of computable a posteriori error bounds. By combining the elliptic reconstruction in space and the three-point reconstruction in time, we rigorously derive a residual-based a posteriori error bound for the second order BDF scheme. This bound serves as the foundation for reliable error estimation as well as the adaptive control of the spatial grid and temporal time-step sizes.

Another focus lies on the stability analysis in the sense of a discrete energy inequality for the fully discrete scheme. We specifically investigate variable time-step sizes as well as the dependence of the discrete energy inequality on the damping constant and derive an explicit connection between the roots of the BDF polynomials and G-stability. This relation is crucial for proving the discrete energy inequality. By utilizing discrete orthogonal convolution kernels corresponding to the BDF coefficients, we further extend the discrete energy inequality to an $\mathcal{L}^2$-bound for the discrete solution.

Building upon these theoretical results, we construct a time integration algorithm with variable time-step sizes and variable orders (VSVO), which is combined with a residual-based error estimator for spatial coarsening and refinement. Finally, we demonstrate the effectiveness of the space and time adaptive algorithm through a variety of numerical examples.

Kurzfassung

Die Landau–Lifshitz–Gilbert Gleichung (LLG) dient als grundlegendes Modell zur Beschreibung der Magnetisierungsdynamik in ferromagnetischen Materialien. Aufgrund ihrer stark nichtlinearen Struktur, die unter anderem eine nicht-konvexe Nebenbedingung zur Erhaltung der Magnetisierungslänge in jedem Punkt in Raum und Zeit umfasst, stellt sie hohe Herausforderungen an die numerische Analysis. Insbesondere kleine Dämpfungskonstanten erschweren die mathematische Stabilitätsuntersuchung und die effiziente numerische Simulation beträchtlich. Da in der Praxis zudem stark lokalisierte Phänomene wie Domänenwände (engl. *domain walls*) im Raum oder schnelle Schaltvorgänge in der Zeit auftreten, liegt der Einsatz von adaptiven Verfahren nahe, um diese lokalen Strukturen effizient aufzulösen.

In dieser Arbeit untersuchen wir ein Tangentialraumverfahren (engl. *tangent plane scheme*) zur Lösung der LLG Gleichung, welches die intrinsische Normierung der Magnetisierung in einem schwachen Sinne berücksichtigt. Wir verwenden dabei Finite Elemente zur räumlichen Diskretisierung und nutzen linear-implizite Formeln zur Rückwärtsdifferenzierung (engl. *backward differentiation formulas* (BDF)) zur Zeitintegration. Während die Konvergenz dieses Verfahrens für höhere Ordnungen bereits bekannt ist, liegt der Schwerpunkt dieser Thesis auf der Untersuchung und Herleitung von berechenbaren a posteriori Fehlerschranken. Durch die Kombination einer elliptischen Rekonstruktion im Raum mit einer Drei-Punkt-Rekonstruktion in der Zeit leiten wir rigoros eine residualbasierte Fehlerschranke für das BDF Verfahren zweiter Ordnung her. Diese Schranke dient als Grundlage für eine zuverlässige Fehlerschätzung sowie für die adaptive Steuerung des räumlichen Gitters und der zeitlichen Schrittweiten.

Ein weiterer Fokus liegt auf der Stabilitätsanalyse im Sinne einer diskreten Energieungleichung für das volldiskrete Verfahren. Wir untersuchen dabei insbesondere variable Zeitschrittweiten sowie die Abhängigkeit von der Dämpfungskonstante und leiten einen expliziten Zusammenhang zwischen den Nullstellen der BDF-Polynome und der G-Stabilität her. Dieser Zusammenhang ist entscheidend für den Nachweis der diskreten Energieungleichung. Unter Verwendung diskreter orthogonaler Faltungskerne (engl. *discrete orthogonal convolution kernels*) zu den BDF Koeffizienten erweitern wir die diskrete Energieungleichung zudem zu einer $\mathcal{L}^2$-Schranke für die diskrete Lösung.

Aufbauend auf den theoretischen Ergebnissen konstruieren wir einen Algorithmus zur Zeitintegration mit variablen Zeitschrittweiten und variablen Ordnungen (VSVO), den wir mit einem residualbasierten Fehlerschätzer zur räumlichen Vergröberung und Verfeinerung kombinieren. Anhand verschiedener numerischer Beispiele demonstrieren wir abschließend die Effektivität des raum- und zeitadaptiven Gesamtverfahrens.

Contents

CHAPTER 1

Introduction

In this chapter, we first provide a brief motivation for the applications of the Landau–Lifshitz–Gilbert (LLG) equation before describing the underlying physical forces and energy contributions. This is followed by the derivation of a dimensionless form of the equation that serves as the basis for our subsequent analysis. We study various mathematically equivalent representations of the LLG equation, including both the Gilbert and the Landau–Lifshitz forms. To provide the necessary context, we present a comprehensive review of the current literature and highlight recent advancements in analytical results and numerical methodologies. Then, we review the analytical foundations of the LLG equation, focusing specifically on the existence and (non-)uniqueness of strong and weak solutions. Finally, we establish our mathematical notation and introduce the fully discrete tangent plane scheme, which provides the fundamental numerical framework for this work.

1.1 Motivation

The physical understanding of magnetization dynamics is essential for a broad range of modern technological applications. This includes semiconductor-based spintronics [124], as well as advanced memory architectures such as spin-orbit torque magnetic random-access memory [68, 85]. The foundation for these technologies was established with the discovery of the giant magnetoresistance effect [20, 34], for which *A. Fert* and *P. Grünberg* were awarded the Nobel Prize of Physics in 2007. This phenomenon enabled efficient control of the motion of electrons through relative magnetic orientation in thin-film layers [69]. This breakthrough led to a remarkable increase in the storage density of digital information, primarily through the development of read heads on hard disk drives.

More recently, magnet skyrmions have emerged as promising information carriers in nanoscale dimension [66, 67]. The development of these next-generation topological storage devices is based on a deep understanding of magnetization dynamics. The fundamental mathematical framework for describing these complex phenomena is the LLG equation.

1.2 Landau–Lifshitz–Gilbert equation

Let $\Omega \subset \mathbb{R}^3$ denote the volume of a ferromagnetic body and let $T > 0$ be a fixed finite time. The magnetic state of the ferromagnet is described by its magnetization, represented by a three-dimensional vector field $\boldsymbol{M} : [0,T] \times \Omega \to \mathbb{R}^3$ (in A/m). Assuming the temperature of the material remains constant and sufficiently low, the magnitude of the magnetization is

preserved, i.e.,

$$|\boldsymbol{M}| = M_s,$$

where M_s is the saturation magnetization. Typically, M_s is of the order $10^5\,\mathrm{A/m}$ and provides a measure of the maximum magnetic field that the material can generate. In static micromagnetism, the main idea is that the magnetization minimizes the total magnetic Gibbs free energy of a ferromagnetic body. The dynamical behavior of the magnetization is described by the LLG equation. This equation describes the precessional motion of the magnetization around the effective field, which is defined as the negative functional derivative of the total magnetic Gibbs free energy. In addition, the LLG equation incorporates a damping term to account for the dissipative effects observed experimentally. In this section, we first introduce various energy contributions considered in this work, which together define the total magnetic energy, before presenting the LLG equation itself. This section is based on [38, 76, 87, 107].

Exchange energy

A fundamental property of a ferromagnet is its tendency to align the magnetization of local regions in the same direction. This alignment results from the exchange interaction, whose contribution to the total energy is referred to as the exchange energy and is defined by

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

where $A > 0$ is the exchange stiffness constant (in $\mathrm{J/m}$). The exchange energy is minimized when adjacent magnetic moments are parallel to each other.

External field energy

The external field energy, often referred to as Zeeman's energy, tends to align the magnetization $\boldsymbol{M}$ in the direction of the external field $\boldsymbol{H}_{\mathrm{ext}}$ and is given by

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

where $\mu_0 = 4\pi \cdot 10^{-7}\,\mathrm{N/A^2}$ describes the permeability of vacuum.

Magnetostatic energy

The magnetization itself induces a magnetic field $\boldsymbol{H} : \mathbb{R}^3 \to \mathbb{R}^3$, which is described by the magnetostatic Maxwell equations

$$\boldsymbol{\nabla} \times \boldsymbol{H} = \boldsymbol{J} \qquad \text{and} \qquad \boldsymbol{\nabla} \cdot \boldsymbol{H} = \boldsymbol{\nabla} \cdot (\chi_\Omega \boldsymbol{M}), \tag{1.2.1}$$

where $\boldsymbol{J}$ is the current density (in $\mathrm{A/m^2}$). The magnetic field, which is generated by the current density, is known as the Oersted field. For further details, we refer to [38, 107]. In this

work, however, we focus on the magnetic field arising from the magnetization itself, commonly referred to as the stray field or demagnetizing field $\boldsymbol{H}_\mathrm{s}$. Since the stray field is an irrotational field, it can be expressed as the gradient of a scalar potential U, i.e., $\boldsymbol{H}_\mathrm{s}(\boldsymbol{M}) = -\boldsymbol{\nabla} U(\boldsymbol{M})$, where the potential $U(\boldsymbol{M})$ depends linearly on the magnetization $\boldsymbol{M}$. Adding suitable transmission conditions and a radiation condition, the scalar potential solves the full-space transmission problem

$$\begin{aligned} -\Delta U^\mathrm{int} &= -\boldsymbol{\nabla}\cdot\boldsymbol{M}, && \text{in } \Omega, \\ -\Delta U^\mathrm{ext} &= 0, && \text{in } \mathbb{R}^3 \setminus \overline{\Omega}, \\ U^\mathrm{ext} - U^\mathrm{int} &= 0, && \text{on } \partial\Omega, \\ \boldsymbol{\nabla}(U^\mathrm{ext} - U^\mathrm{int})\cdot\boldsymbol{n} &= -\boldsymbol{M}\cdot\boldsymbol{n}, && \text{on } \partial\Omega, \\ U(\boldsymbol{x}) &= \mathcal{O}(1/|\boldsymbol{x}|), && \text{for } |\boldsymbol{x}| \to \infty. \end{aligned} \tag{1.2.2}$$

Finally, the energy contribution of the stray field is given by

$$\mathcal{E}_\mathrm{stray}(\boldsymbol{M}) = -\mu_0 \int_\Omega \boldsymbol{H}_\mathrm{s}(\boldsymbol{M})\cdot\boldsymbol{M}\,\mathrm{d}\boldsymbol{x}.$$

Total magnetic Gibbs free energy

By summing the contributions outlined above, we obtain the total magnetic Gibbs free energy of a ferromagnetic body by

$$\mathcal{E}_\mathrm{tot}(\boldsymbol{M}) = \mathcal{E}_\mathrm{ex}(\boldsymbol{M}) + \mathcal{E}_\mathrm{ext}(\boldsymbol{M}) + \mathcal{E}_\mathrm{stray}(\boldsymbol{M}). \tag{1.2.3}$$

Additional contributions may arise, for instance, from magneto-elastic effects, Dzyaloshinskii–Moriya interactions, magnetostriction, anisotropy, or the Oersted field [38, 87, 107]. Following the theory of micromagnetism, the magnetization of a ferromagnet is obtained by minimizing the total Gibbs free energy

$$\min_{\boldsymbol{M}\,:\,|\boldsymbol{M}|=M_s} \mathcal{E}_\mathrm{tot}(\boldsymbol{M}).$$

The effective field is defined as the negative functional derivative of the Gibbs free energy. Thus, we obtain by variation

$$\mu_0 \boldsymbol{H}_\mathrm{eff}(\boldsymbol{M}) = -\frac{\delta \mathcal{E}_\mathrm{tot}}{\delta \boldsymbol{M}}(\boldsymbol{M}) = \frac{2A}{M_s^2}\Delta\boldsymbol{M} + \mu_0 \boldsymbol{H}_\mathrm{ext} + \mu_0 \boldsymbol{H}_\mathrm{s}(\boldsymbol{M}). \tag{1.2.4}$$

The origin of the LLG equation dates back to [95], published in 1935, in which Lifshitz and Landau introduced the following equation to model the dynamics of the magnetization driven by the effective field

$$\partial_t \boldsymbol{M} = -\frac{\gamma_0}{1+\alpha^2}\boldsymbol{M}\times\boldsymbol{H}_\mathrm{eff}(\boldsymbol{M}) - \frac{\alpha\gamma_0}{(1+\alpha^2)M_s}\boldsymbol{M}\times(\boldsymbol{M}\times\boldsymbol{H}_\mathrm{eff}(\boldsymbol{M})), \tag{1.2.5}$$

where the variable $\gamma_0 \approx 2.21\cdot 10^5\,\mathrm{m/(A\,s)}$ describes the gyromagnetic ratio. Equation

(1.2.5) consists of a precessional term $\boldsymbol{M} \times \boldsymbol{H}_{\text{eff}}(\boldsymbol{M})$ that captures the rotational motion of the magnetization vector $\boldsymbol{M}$ around the effective field $\boldsymbol{H}_{\text{eff}}(\boldsymbol{M})$ and a damping term $\boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}_{\text{eff}}(\boldsymbol{M}))$ that drives the relaxation of the magnetization $\boldsymbol{M}$ toward the effective field $\boldsymbol{H}_{\text{eff}}(\boldsymbol{M})$. The dimensionless parameter $\alpha > 0$ acts as an empirical constant that controls the intensity of the damping. In his PhD thesis [74, 75], Gilbert proposed a different damping term

$$\partial_t \boldsymbol{M} = -\gamma_0 \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}}(\boldsymbol{M}) + \frac{\alpha}{M_s} \boldsymbol{M} \times \partial_t \boldsymbol{M}. \tag{1.2.6}$$

However, as we see in the following section, these formulations are mathematically equivalent. The Gilbert form (1.2.6) is frequently referred to as the Landau–Lifshitz–Gilbert (LLG) equation in the literature.

Dimensionless Landau–Lifshitz–Gilbert equation

We now derive the dimensionless form of the LLG equation and introduce equivalent formulations. A non-dimensional derivation can be found, e.g., in [38]. In general, we denote the corresponding dimensionless variables with small letters. Similar proofs for equivalent formulations are given in [7, 76].

We start by defining dimensionless variables $t' \coloneqq T_{\text{ref}} t$ and $\boldsymbol{x}' \coloneqq L_{\text{ref}} \boldsymbol{x}$. A common choice for the time scale, which we employ, is given by $T_{\text{ref}} = \gamma_0 M_s$. A possible choice for the reference length is given by $L_{\text{ref}} \coloneqq \sqrt{2A/(\mu_0 M_s^2)}$, but we keep L_{ref} in its general form. We further introduce the dimensionless magnetization

$$\boldsymbol{m}(t', \boldsymbol{x}') \coloneqq \frac{1}{M_s} \boldsymbol{M}\Big(\frac{t'}{\gamma_0 M_s}, \frac{\boldsymbol{x}'}{L_{\text{ref}}}\Big), \qquad (t', \boldsymbol{x}') \in \Omega_T', \tag{1.2.7}$$

where Ω_T' denotes the time-space domain obtained from $\Omega_T \coloneqq (0, T) \times \Omega$ by rescaling the time with T_{ref} and the space with L_{ref}. For notational simplicity, in the remainder of this work, we omit the primes and write the dimensionless time and space as $(t, \boldsymbol{x})$ and the corresponding domain as Ω_T, with the meaning being clear from the context. Defining the dimensionless effective field $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) \coloneqq \boldsymbol{H}_{\text{eff}}(\boldsymbol{M})/M_s$, as well as the dimensionless external field $\boldsymbol{h}_{\text{ext}} \coloneqq \boldsymbol{H}_{\text{ext}}/M_s$ and the dimensionless stray field $\boldsymbol{h}_{\text{s}}(\boldsymbol{m}) \coloneqq \boldsymbol{H}_{\text{s}}(\boldsymbol{M})/M_s$, we obtain with (1.2.4)

$$\begin{aligned} \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) &= \frac{2A}{\mu_0 M_s^2 L_{\text{ref}}^2} \Delta \boldsymbol{m} + \frac{1}{M_s} \boldsymbol{H}_{\text{ext}} + \frac{1}{M_s} \boldsymbol{H}_{\text{s}}(\boldsymbol{M}) \\ &= \ell_{\text{ex}}^2 \Delta \boldsymbol{m} + \boldsymbol{h}_{\text{ext}} + \boldsymbol{h}_{\text{s}}(\boldsymbol{m}), \end{aligned} \tag{1.2.8}$$

where the exchange length is defined by

$$\ell_{\text{ex}}^2 \coloneqq \frac{2A}{\mu_0 M_s^2 L_{\text{ref}}^2}. \tag{1.2.9}$$

Starting from the Landau–Lifshitz equation (1.2.5), we derive the dimensionless Landau–

Lifshitz equation for $\boldsymbol{m} : [0,T] \times \Omega \to \mathbb{R}^3$ utilizing (1.2.8) by

$$\partial_t \boldsymbol{m} = -\frac{1}{1+\alpha^2}\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - \frac{\alpha}{1+\alpha^2}\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})). \tag{1.2.10}$$

On the other hand, starting from the Gilbert form (1.2.6) yields the dimensionless Gilbert equation

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) + \alpha \boldsymbol{m} \times \partial_t \boldsymbol{m}. \tag{1.2.11}$$

Since $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{a} = 0$ for $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^3$, multiplying (1.2.10) and (1.2.11) by $\boldsymbol{m}$ immediately yields $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$. Thus, considering the initial condition $\boldsymbol{m}_0(\boldsymbol{x}) \coloneqq \boldsymbol{m}(0, \boldsymbol{x})$ with $|\boldsymbol{m}_0(\boldsymbol{x})| = 1$ a.e. for $\boldsymbol{x} \in \Omega$ and $\frac{1}{2}\partial_t |\boldsymbol{m}|^2 = \partial_t \boldsymbol{m} \cdot \boldsymbol{m} = 0$, we obtain $|\boldsymbol{m}| = 1$ a.e. in Ω_T.

We now proceed by establishing the equivalence between the Landau–Lifshitz equation (1.2.10) and the Gilbert equation (1.2.11). Taking the cross product of $\boldsymbol{m}$ with (1.2.11) yields

$$\boldsymbol{m} \times \partial_t \boldsymbol{m} = -\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})) + \alpha \boldsymbol{m} \times (\boldsymbol{m} \times \partial_t \boldsymbol{m}). \tag{1.2.12}$$

Inserting (1.2.12) into the Gilbert form (1.2.11), we obtain

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - \alpha \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})) + \alpha^2 \boldsymbol{m} \times (\boldsymbol{m} \times \partial_t \boldsymbol{m}). \tag{1.2.13}$$

Using the vector identity

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c} \tag{1.2.14}$$

for $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$, $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$ and $|\boldsymbol{m}| = 1$ a.e. in Ω_T, we obtain

$$\boldsymbol{m} \times (\boldsymbol{m} \times \partial_t \boldsymbol{m}) = (\boldsymbol{m} \cdot \partial_t \boldsymbol{m})\boldsymbol{m} - |\boldsymbol{m}|^2 \partial_t \boldsymbol{m} = -\partial_t \boldsymbol{m}. \tag{1.2.15}$$

Finally, we conclude the equivalence of the Gilbert and the Landau–Lifshitz form by rewriting the right-hand side of (1.2.13) using (1.2.15). We proceed by deriving an alternative form of the LLG equation which serves as the starting point for our analysis. We start by taking the cross product of $\boldsymbol{m}$ with (1.2.10) and adding α times (1.2.10) to obtain

$$\begin{aligned}
\alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} &= -\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})) - \frac{\alpha}{1+\alpha^2}\boldsymbol{m} \times (\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}))) \\
&\quad - \frac{\alpha}{1+\alpha^2}\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) \\
&= -\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})),
\end{aligned}$$

since $\boldsymbol{m} \times (\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}))) = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})$. Using the vector identity (1.2.14) again, we get

$$-\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m})) = \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - (\boldsymbol{m} \cdot \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}))\boldsymbol{m}.$$

Finally, by combining the above equations, we obtain the alternative form

$$\alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) - (\boldsymbol{m} \cdot \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}))\boldsymbol{m}. \tag{1.2.16}$$

In contrast to the Landau–Lifshitz (1.2.10) and the Gilbert formulations (1.2.11), we have to explicitly impose the non-convex constraint $|\boldsymbol{m}| = 1$ a.e. in Ω_T for the alternative form (1.2.16), since the condition $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$ does not follow from the equation itself [76, Remark 1.2.3]. For both the Landau–Lifshitz (1.2.10) and the Gilbert formulation (1.2.11), the initial condition $|\boldsymbol{m}_0| = 1$ a.e. in Ω is sufficient to preserve the unit-length constraint $|\boldsymbol{m}| = 1$ a.e. in Ω_T.

1.3 STATE OF THE ART

The Landau–Lifshitz–Gilbert (LLG) equation provides a fundamental description of the time evolution of the magnetization in a ferromagnetic material. It captures the precessional motion of magnetic moments around the effective field together with a dissipative damping motion [75, 95, 106]. A characteristic physical phenomenon of ferromagnetic materials is the formation of domain walls, sharp interfaces separating nearly uniformly magnetized domains [112], see Fig. 1.3.1. These highly localized phenomena can be employed, for instance, in domain wall based racetrack memory, where the information is stored by pairs, each consisting of a domain wall and a magnetic domain [66]. Recent developments, such as skyrmion-based racetrack memory and spin-orbit torque magnetic random-access memories, focus on another way to store data magnetically [68, 85]. Similarly, the temporal behavior is characterized by rapid changes stemming from switching processes that occur within relatively stable periods. Consequently, the inefficiency of nonadaptive simulations leads to a substantial demand for adaptive refinement in both space and time to resolve the underlying physical and magnetic features.

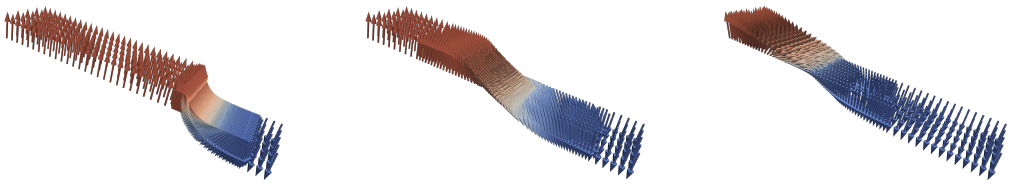


Figure 1.3.1: Illustration of a domain wall, a sharp interface separating different magnetized domains, moving along the magnetic domain. An adaptive algorithm refines and coarsens the spatial grid around the domain wall.

First results regarding the existence of weak solutions of the LLG equation, which are global in time, are given in [14, 79, 120]. However, in general, weak solutions are not unique. For instance, in [14], the authors construct infinitely many weak solutions for $\boldsymbol{m}_0 \in \boldsymbol{\mathcal{H}}^1(\Omega)$ based on ideas in [33, 47] using nonstationary (weakly) harmonic maps. On the other hand, if the initial data is sufficiently smooth, the solution is unique in a strong sense up to some finite time $T^\star$, or global in time if the initial data is of small energy (in 2D) [40, 41]. Numerical methods, in particular those of higher order as considered in this thesis, generally require increased

spatial and temporal regularity of the solution to guarantee convergence. Accordingly, the existence of unique (nontrivial) strong solutions with arbitrary regularity up to any given stopping time has been established in [64].

Numerical methods for the LLG equation have been extensively studied in the literature. For an overview, we refer to the review articles [45, 89]. Starting from the alternative form (1.2.16) of the LLG equation, Alouges and Jaisson developed in [13] the so called tangent plane scheme, where they define $\boldsymbol{v} = \partial_t \boldsymbol{m}$ and solve for $\boldsymbol{v}$ such that (1.2.16) and $\boldsymbol{v} \cdot \boldsymbol{m} = 0$ are satisfied. The authors proposed an explicit time integration scheme and proved weak convergence to a weak global solution, provided that first the time-step size τ tends to 0, followed by the mesh size h. This convergence result was later improved in [29] by requiring $\tau/h^{1+d/2} \to 0$. Moreover, in [16] the method was extended to an implicit finite element scheme. It is important to mention that the scheme in [13] included a projection step onto the unit sphere, whereby the finite element approximation of the magnetization at each node, $\boldsymbol{m}_i$, is normalized so that $|\mathbf{m}_i| = 1$ holds. The error analysis of the scheme from [13] was recently provided in [17] and requires the lower and upper time-step restrictions $\mathcal{O}(h^{7/4}) \leq \tau \leq \mathcal{O}(h)$ due to the difficulties arising from the projection (normalization) step. Although the projection step is well defined, it requires restrictive conditions in fully discrete situations to guarantee stability [31]. On the other hand, the omission of the nodal projection may result in a phase shift in the evolution of $\boldsymbol{m}$ [86]. In recent years, the analysis of projection-free methods, i.e., schemes excluding the normalization step at each node, has attracted increasing attention. In [31], Bartels proposed a predictor-corrector approach for the harmonic map heat flow in which the projection step can be omitted and the normalization constraint is controlled linearly by the time-step size. Such a scheme was analyzed in [1] for the LLG equation while an error analysis is provided in [65] for the eddy current equations coupled to the LLG equation. In [7], the authors introduced a tangent plane scheme based on linearly implicit backward differentiation formulas (BDF) combined with a non-conforming finite element spatial discretization, where the normalization constraint is enforced in an $\mathcal{L}^2$-averaged sense over the finite element basis functions. Furthermore, Akrivis et al. established optimal a priori error estimates of higher order in both space and time for sufficiently regular solutions, along with a discrete energy inequality. Moreover, they consider the pointwise tangent space and observe that, although the same technique achieves optimal-order convergence, the corresponding energy estimate does not admit an h- and τ-independent bound that is irrespective of the smoothness of the solution [7, Remark 2.2]. A related approach based on the BDF scheme with a nodal treatment of the normalization constraint, which is first order in time and space, is given in [30]. More recently, [10] introduced a projection-free method that achieves quadratically accurate enforcement of the normalization constraint with respect to the (uniform) time-step size using BDF(2). In [11], the authors proposed a linearly implicit (θ, μ)-method which, while ensuring second-order accuracy of the normalization constraint similar to [10], exhibits improved stability properties compared to the BDF scheme under variable time-step size strategies.

The presence of domain walls, which are sharp interfaces that separate regions with nearly

uniform magnetization, within a magnetic domain creates a strong demand for adaptive methods that can refine the computational mesh for the LLG equations. Various adaptive strategies have been explored in the literature: moving mesh finite element methods that dynamically re-distribute nodes to regions where the magnetization exhibits a large gradient [63] or adaptive mesh refinement on finite difference grids in 2D with refinement indicators such as the in-plane divergence of the magnetization, the gradient of the angle formed by the in-plane components of the magnetization and the norm of the rotational field [72]. In [19], the violation of the normalization constraint is utilized to adapt the spatial mesh for the static micromagnetic equations. Further computational applications of adaptive mesh refinement techniques appear in [84, 111, 116]. On the other hand, a variety of adaptive time-stepping strategies have been investigated for the LLG equation. For example, [114] proposes an adaptive scheme based on the implicit midpoint rule, where the error is estimated using an explicit BDF(3) method. Applications of adaptive time-stepping schemes based on the extrapolation technique of [81] can be found in [62, 76]. Furthermore, adaptive time-stepping for the harmonic map heat flow using a linearly implicit (θ, μ)-method is discussed in [11].

However, the mathematical analysis of a posteriori error estimators for the LLG equation remains limited. A notable contribution is given in [23], where a posteriori estimates are established for both a space-time discretization and a semi-implicit scheme, formulated without enforcing the normalization constraint $|\boldsymbol{m}| = 1$. While the semi-implicit method is structurally related to the approach in [106], it inherits stability issues due to the unconstrained formulation. Related a posteriori results for a LLG model including magnetostriction, but without the exchange field, are provided in [21, 24].

1.4 NOTATION

In this section, we establish the mathematical notation utilized throughout this work. For a more comprehensive introduction, we refer to [2, 61, 98]. For the convenience of the reader, a concise summary of the most frequently used symbols is provided in the Abbreviations and Symbols section at the end of the thesis.

For a vector $v \in \mathbb{C}^n$, we denote the maximum norm by

$$\|v\|_\infty \coloneqq \max_{i=1,\dots,n} |v_i| \,. \tag{1.4.1}$$

Further, for $A = (a_{ij}) \in \mathbb{C}^{m\times n}$ we define the induced ℓ^∞-matrix norm, also known as the maximum absolute row sum norm, as

$$\|A\|_\infty \coloneqq \max_{x\in\mathbb{C}^n \,:\, \|x\|_\infty = 1} \|Ax\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}|. \tag{1.4.2}$$

In general, we denote the identity matrix in $\mathbb{R}^{d\times d}$ by I_d.

Function spaces on Ω

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded domain with Lipschitz boundary. We denote the standard Lebesgue spaces by $\mathcal{L}^p(\Omega)$ and usual Sobolev spaces by $\mathcal{W}^{k,p}(\Omega)$ for $1 \leq p \leq \infty$ and $k \in \mathbb{N}$. The Sobolev spaces are equipped with the norms $\|\cdot\|_{\mathcal{W}^{k,p}(\Omega)}$ and semi-norms $|\cdot|_{\mathcal{W}^{k,p}(\Omega)}$. We employ the common short-hand notation $\mathcal{H}^k(\Omega) \coloneqq \mathcal{W}^{k,2}(\Omega)$. Additionally, we denote the standard $\mathcal{L}^2$-inner product and its associated norm by $(u, v) = \int_\Omega uv \,\mathrm{d}x$ and $\|u\| = \|u\|_{\mathcal{L}^2(\Omega)} = \sqrt{(u, u)}$. We omit the subscript of a norm whenever its domain is clear from the context. If X is a normed space, its dual X^* consists of all continuous linear functions $g : X \to \mathbb{C}$. We denote the duality pairing between $g \in X^*$ and $u \in X$ by $\langle g, u\rangle = g(x)$. If $g, u \in \mathcal{L}^2(\Omega)$, the duality pairing $\langle g, u\rangle$ coincides with the $\mathcal{L}^2$-inner product (g, u), i.e., $\langle g, u\rangle = (g, u)$.

Throughout this work, bold symbols represent vector-valued quantities. Specifically, we define the vector-valued Sobolev spaces $\boldsymbol{\mathcal{W}}^{k,p}(\Omega) \coloneqq (\mathcal{W}^{k,p}(\Omega))^3$ and $\boldsymbol{\mathcal{H}}^k(\Omega) \coloneqq (\mathcal{H}^k(\Omega))^3$. This convention extends analogously to other function spaces. The space of infinitely differentiable functions with compact support in Ω is denoted by $\mathcal{C}_c^\infty(\Omega)$, and $\mathcal{H}_0^1(\Omega)$ represents its closure under the $\mathcal{H}^1$-norm. The dual space of $\mathcal{H}_0^1(\Omega)$ is denoted by $\mathcal{H}^{-1}(\Omega)$, and the dual of $\mathcal{H}^1(\Omega)$ is denoted by $\widetilde{\mathcal{H}}^{-1}(\Omega)$. Furthermore, we introduce the space

$$\mathcal{H}_\star^1(\Omega) \coloneqq \{v \in \mathcal{H}^1(\Omega) \,:\, \Delta v \in \mathcal{L}^2(\Omega) \text{ and } \partial_{\boldsymbol{n}} v = 0 \text{ on } \partial\Omega\}.$$

Moreover, we define the Slobodeckij semi-norm

$$|u|^2_{\mathcal{H}^{1/2}(\partial\Omega)} = \int_{\partial\Omega}\int_{\partial\Omega} \frac{|v(\boldsymbol{x}) - v(\boldsymbol{y})|^2}{|\boldsymbol{x} - \boldsymbol{y}|^d} \,\mathrm{d}\sigma_{\boldsymbol{y}} \,\mathrm{d}\sigma_{\boldsymbol{x}}.$$

This allows us to conveniently define the fractional Sobolev space

$$\mathcal{H}^{1/2}(\partial\Omega) \coloneqq \Big\{u \in \mathcal{L}^2(\partial\Omega) \;:\; |u|_{\mathcal{H}^{1/2}(\partial\Omega)} < \infty\Big\},$$

equipped with the norm $\|u\|^2_{\mathcal{H}^{1/2}(\partial\Omega)} \coloneqq \|u\|^2_{\mathcal{L}^2(\partial\Omega)} + |u|^2_{\mathcal{H}^{1/2}(\partial\Omega)}$. Finally, we define the dual space of $\mathcal{H}^{1/2}(\partial\Omega)$ by $\mathcal{H}^{-1/2}(\partial\Omega) \coloneqq (\mathcal{H}^{1/2}(\partial\Omega))^*$.

Bochner spaces

Let X be a Banach space and $T > 0$. For $1 \leq p \leq \infty$, the Bochner–Lebesgue spaces are defined by

$$\mathcal{L}^p(0, T; X) \coloneqq \{v : (0, T) \to X \,:\, v \text{ is measurable and } \|v\|_{\mathcal{L}^p(0,T;X)} < \infty\}$$

with the corresponding norm

$$\|v\|_{\mathcal{L}^p(0,T;X)} = \begin{cases} \Big(\int_0^T \|v(t)\|_X^p \,\mathrm{d}t\Big)^{1/p}, & \text{for } 1 \leq p < \infty, \\ \operatorname*{ess\,sup}_{t\in(0,T)} \|v(t)\|_X, & \text{for } p = \infty. \end{cases}$$

Furthermore, we define the space of k-times weakly differentiable functions in $\mathcal{L}^2(0,T;X)$ by

$$\mathcal{H}^k(0,T;X) \coloneqq \{v \in \mathcal{L}^2(0,T;X) \,:\, \partial_t^\ell v \in \mathcal{L}^2(0,T;X) \quad \text{for } 1 \leq \ell \leq k\}$$

for $k \in \mathbb{N}$. Moreover, we define with $\Omega_T = (0,T) \times \Omega$ the space

$$\mathcal{H}^{i,2i}(\Omega_T) \coloneqq \{v \in \mathcal{L}^2(0,T;\mathcal{L}^2(\Omega)) \,:\, \|v\|_{\mathcal{H}^{i,2i}(\Omega_T)} < \infty\}$$

with the corresponding norm

$$\|v\|_{\mathcal{H}^{i,2i}(\Omega_T)} \coloneqq \sum_{\ell=0}^{i} \|\partial_t^\ell v\|_{\mathcal{L}^2(0,T;\mathcal{H}^{2i-2\ell}(\Omega))}$$

for $i \in \mathbb{N}_0$.

Discretization and trace operators

We denote by $\{\mathcal{T}_h^n\}_{n=0}^N$ a family of shape regular triangulations of the domain Ω. Additionally, let Σ_n represent the set of all edges (for $d=2$) or faces (for $d=3$) of $\mathcal{T}_h^n$ and denote by Σ_n° the set of all interior edges or faces. The space of globally continuous piecewise polynomials of degree p on $\mathcal{T}_h^n$ is given by

$$\mathcal{S}_h^p(\mathcal{T}_h^n) \coloneqq \{u \in C^0(\overline{\Omega}) \,:\, u|_K \in \mathbb{P}_p \quad \text{for all } K \in \mathcal{T}_h^n\},$$

where $\mathbb{P}_p$ refers to the space of polynomials of degree p. For any $E \in \Sigma_n^\circ$ with associated normal vector $\boldsymbol{n}_E$, the jump across the interface E is defined by

$$[\![v_h]\!]_E(\boldsymbol{x}) \coloneqq \lim_{s \to 0^+} \left(v_h(\boldsymbol{x} + s\boldsymbol{n}_E) - v_h(\boldsymbol{x} - s\boldsymbol{n}_E)\right).$$

Moreover, we introduce appropriate trace operators to characterize the behavior of functions at the boundary of the domain Ω. The following theorem from [98, Theorem 3.37] defines the Dirichlet trace operator.

Theorem 1 (Dirichlet trace operator)**.** *Let $\Omega \subset \mathbb{R}^d$, $d = 2,3$, be a bounded domain with Lipschitz boundary. Then there exists a unique operator $\gamma_D : \mathcal{H}^1(\Omega) \to \mathcal{H}^{1/2}(\partial\Omega)$ with*

$$\gamma_D u = u|_{\partial\Omega}$$

for all $u \in C^0(\overline{\Omega})$.

A generalization of this result to the spaces $H^s(\Omega)$ for $s > 1/2$ is also given in [98, Theorem 3.37]. To describe the normal derivative on the boundary in a weak sense, we utilize the Neumann trace operator γ_N characterized by the following lemma from [98, Lemma 4.3].

Lemma 2 (Neumann trace operator)**.** *Let $\Omega \subset \mathbb{R}^d$, $d = 2,3$, be a bounded domain with*

Lipschitz boundary. Then, there exists a unique operator $\gamma_N : \mathcal{H}^1(\Omega) \mapsto \mathcal{H}^{-1/2}(\partial\Omega)$ *with*

$$\langle \boldsymbol{\nabla} u, \boldsymbol{\nabla}\phi\rangle = \langle f, \phi\rangle + \langle \gamma_N u, \gamma_D \phi\rangle, \qquad \forall \phi \in \mathcal{H}^1(\Omega),$$

if $u \in \mathcal{H}^1(\Omega)$ *and* $f \in \widetilde{\mathcal{H}}^{-1}(\Omega)$ *satisfy* $-\Delta u = f$ *in a distributional sense.*

In particular, if $u \in \mathcal{C}^2(\overline{\Omega})$, then $\gamma_{\mathrm{N}} u = \partial_{\boldsymbol{n}} u = \boldsymbol{\nabla} u \cdot \boldsymbol{n}$, where $\boldsymbol{n}$ denotes the outer normal vector of Ω.

1.5 Weak and strong solutions of the LLG equation

In this section, we recall the basic analytical framework for the LLG equation. We introduce the notions of weak and strong solutions for a bounded domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, and summarize classical results on the existence and uniqueness of strong solutions, together with the existence and non-uniqueness of weak solutions. Finally, we draw a connection between weak and strong solutions with the weak-strong uniqueness principle.

We begin by introducing the notion of a strong solution. To this end, we first recall the alternative (dimensionless) form of the LLG equation (1.2.16) by

$$\alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) - (\boldsymbol{m} \cdot \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}))\boldsymbol{m} = \boldsymbol{P}(\boldsymbol{m})\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}), \tag{1.5.1}$$

where $\boldsymbol{P}(\boldsymbol{m}) = \boldsymbol{I} - \boldsymbol{m}\boldsymbol{m}^{\top}$ denotes the orthogonal projection onto the tangent plane to the unit sphere $\mathbb{S}^2$ at $\boldsymbol{m}$, provided that $|\boldsymbol{m}| = 1$. Based on this representation, we define strong solutions as follows.

Definition 3 (Strong solution)**.** *A function* $\boldsymbol{m} : [0, T] \times \Omega \to \mathbb{R}^3$ *is called a strong solution of the LLG equation if* $\boldsymbol{m} \in \mathcal{L}^\infty(0, T; \boldsymbol{\mathcal{H}}^2(\Omega)) \cap \mathcal{H}^1(0, T; \boldsymbol{\mathcal{L}}^2(\Omega))$ *satisfies*

$$\begin{cases} \alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = \boldsymbol{P}(\boldsymbol{m})\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}), & \text{in } (0, T) \times \Omega, \\ |\boldsymbol{m}| = 1, & \text{a.e. in } (0, T) \times \Omega, \\ \partial_{\boldsymbol{n}} \boldsymbol{m} = 0, & \text{on } (0, T) \times \partial\Omega, \\ \boldsymbol{m}(0, \cdot) = \boldsymbol{m}_0, & \text{in } \Omega, \end{cases} \tag{1.5.2}$$

where

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

and $\alpha > 0$ *denotes the damping parameter,* $\ell_{\mathrm{ex}} > 0$ *the exchange length and* $\boldsymbol{h}_{\mathrm{ext}}$ *an external field. The initial data* $\boldsymbol{m}_0$ *satisfies* $|\boldsymbol{m}_0| = 1$ *a.e. in* Ω *and* $\partial_{\boldsymbol{n}} \boldsymbol{m}_0 = 0$ *on* $\partial\Omega$*.*

The following theorem ensures local-in-time existence and uniqueness of strong solutions to (1.5.2) for sufficiently regular initial data. This is a classical result of [40, Theorem 1.3]. Moreover, in the two-dimensional case, the solution exists globally in time provided the initial data has sufficiently small energy $\|\boldsymbol{\nabla} \boldsymbol{m}_0\|$, see [40, Theorem 1.4].

Theorem 4 (Existence of strong solutions)**.** *Let $\boldsymbol{h}_{\text{ext}} = \mathbf{0}$. Assume that the initial data $\boldsymbol{m}_0$ satisfies $\boldsymbol{m}_0 \in \boldsymbol{\mathcal{H}}^2(\Omega)$. Then, there exists a time $T^\star > 0$, such that* (1.5.2) *has a unique strong solution $\boldsymbol{m} \in \mathcal{C}^0([0,T];\boldsymbol{\mathcal{H}}^2(\Omega)) \cap \mathcal{L}^2(0,T;\boldsymbol{\mathcal{H}}^3(\Omega))$ for all $T < T^\star$.*

While the previous theorem provides local-in-time well-posedness for strong solutions under minimal regularity assumptions, stronger results can be obtained if the initial data is smoother and satisfies additional conditions. In this case, we can derive (nontrivial) strong solution with higher regularity. The following theorem, provided in [64, Theorem 2.1], establishes the existence and uniqueness of strong solutions with higher regularity.

Theorem 5 (Existence of regular strong solutions)**.** *Let $\boldsymbol{h}_{\text{ext}} = \mathbf{0}$. Assume that the initial data $\boldsymbol{m}_0$ satisfies, for $i \geq 3$,*

$$\boldsymbol{m}_0 \in \boldsymbol{\mathcal{H}}^{2i}(\Omega) \cap \boldsymbol{\mathcal{H}}^1_\star(\Omega), \qquad D^j \boldsymbol{m}_0 \in \boldsymbol{\mathcal{H}}^1_\star(\Omega)$$

for all partial derivatives D^j of order j with $j \leq 2(i-1)$. Then there exists a constant C_i such that $|\boldsymbol{m}_0|_{\boldsymbol{\mathcal{H}}^{2i}(\Omega)} \leq C_i$ implies that (1.5.2) *has a unique strong solution $\boldsymbol{m} \in \boldsymbol{\mathcal{H}}^{i,2i}(\Omega_T)$. The solution $\boldsymbol{m}$ satisfies*

$$\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{H}}^{i,2i}(\Omega_T)} \leq \widetilde{C}_i \|\boldsymbol{m}_0\|_{\boldsymbol{\mathcal{H}}^{2i}}.$$

The constants $C_i, \widetilde{C}_i$ depend only on $\alpha, \Omega, \ell_{\text{ex}}, T$ and i.

Formally, starting from the strong formulation in the Gilbert form (1.2.11), we multiply the equation by appropriate test functions and integrate by parts in space and time to obtain a weak formulation of the LLG equation. Moreover, the definition of a weak solution typically includes an energy inequality, which ensures the physical admissibility of the solution. The following definition of a weak solution follows [13, 14, 57].

Definition 6 (Weak solution)**.** *A function $\boldsymbol{m} : [0,T] \times \Omega \to \mathbb{R}^3$ is called a weak solution of the LLG equation if*

(i) $\boldsymbol{m} \in \mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{H}}^1(\Omega)) \cap \mathcal{H}^1(0,T;\boldsymbol{\mathcal{L}}^2(\Omega))$ *and* $|\boldsymbol{m}| = 1$ *a.e. in* Ω_T*,*

(ii) $\boldsymbol{m}(0,\cdot) = \boldsymbol{m}_0$ *in the sense of traces,*

(iii) for all $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{H}}^1(\Omega_T)$ *it holds that*

$$\int_0^T \langle \partial_t \boldsymbol{m}, \boldsymbol{\varphi} \rangle \,\mathrm{d}t = \int_0^T \Big(\alpha \langle \partial_t \boldsymbol{m}, \boldsymbol{\varphi} \times \boldsymbol{m} \rangle + \ell_{\text{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}, \boldsymbol{\nabla}(\boldsymbol{\varphi} \times \boldsymbol{m}) \rangle + \langle \boldsymbol{h}_{\text{ext}}, \boldsymbol{\varphi} \times \boldsymbol{m} \rangle \Big) \,\mathrm{d}t$$

(iv) the energy inequality

$$\mathcal{E}_{\text{tot}}(\boldsymbol{m}(t)) + \int_0^t \Big(\alpha \|\partial_t \boldsymbol{m}(t)\|^2 + \langle \partial_t \boldsymbol{h}_{\text{ext}}(t), \boldsymbol{m}(t) \rangle \Big) \,\mathrm{d}t \leq \mathcal{E}_{\text{tot}}(\boldsymbol{m}_0)$$

holds for all $t \in (0,T)$, where $\mathcal{E}_{\text{tot}}(\boldsymbol{m}) = \mathcal{E}_{\text{ex}}(\boldsymbol{m}) + \mathcal{E}_{\text{ext}}(\boldsymbol{m})$ with $\mathcal{E}_{\text{ex}}(\boldsymbol{m}) = \frac{\ell_{\text{ex}}^2}{2} \|\boldsymbol{m}\|^2$ $\mathcal{E}_{\text{ext}}(\boldsymbol{m}) = -\langle \boldsymbol{h}_{\text{ext}}, \boldsymbol{m} \rangle$.

The following theorem is a classical result for the existence of weak solutions in the absence of an external field provided in [14, Theorem 1.5].

Theorem 7 (Existence of weak solutions)**.** *Let $\boldsymbol{h}_{\text{ext}} = \mathbf{0}$. Then, for all $\boldsymbol{m}_0 \in \boldsymbol{\mathcal{H}}^1(\Omega)$ with $|\boldsymbol{m}_0| = 1$ a.e. in Ω, there exists a weak solution in the sense of Definition 6.*

Although the previous theorem guarantees the existence of weak solutions, their uniqueness is generally not ensured. By employing nonstationary (weakly) harmonic maps, [14, Theorem 1.6] constructs infinitely many weak solutions, as summarized in the following theorem.

Theorem 8 (Non-uniqueness of weak solutions)**.** *Let $\boldsymbol{h}_{\text{ext}} = \mathbf{0}$. Then there exists $\boldsymbol{m}_0 \in \boldsymbol{\mathcal{H}}^1(\Omega)$ with $|\boldsymbol{m}_0| = 1$ a.e. in Ω, such that one can find infinitely many weak solutions in the sense of Definition 6.*

A connection between weak and strong solutions is established by the weak-strong uniqueness principle given in [57, Theorem 3]. Specifically, if a sufficiently regular strong solution exists, then it coincides with any weak solution originating from the same initial data, as stated in the following theorem.

Theorem 9 (Weak-strong uniqueness principle)**.** *Let $\boldsymbol{m}_0 \in \boldsymbol{\mathcal{H}}^1(\Omega)$. If $\boldsymbol{m}_s \in \mathcal{C}^3(\overline{\Omega}_T)$ is a strong solution of* (1.5.2) *and $\boldsymbol{m}_w$ is a weak solution in the sense of Definition 6. Then*

$$\boldsymbol{m}_s = \boldsymbol{m}_w \qquad \text{a.e. in } \Omega_T.$$

1.6 Tangent plane scheme

In this section, we introduce the fully discrete scheme studied in this thesis. The numerical scheme is based on the formulation of [7, Sct. 2.2]. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded domain and $\boldsymbol{m} : [0, T] \times \Omega \to \mathbb{R}^3$ be a strong solution of (1.5.2).

We start by recalling the alternative form of the LLG equation (1.5.1) with the normalization constraint given by

$$\alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} = \boldsymbol{P}(\boldsymbol{m}) \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}), \tag{1.6.1}$$

$$\partial_t \boldsymbol{m} \cdot \boldsymbol{m} = 0, \tag{1.6.2}$$

where $\boldsymbol{P}(\boldsymbol{m}) = \boldsymbol{I} - \boldsymbol{m}\boldsymbol{m}^\top$ denotes the orthogonal projection onto the tangent plane to the unit sphere $\mathbb{S}^2$ at $\boldsymbol{m}$, provided that $|\boldsymbol{m}| = 1$. We consider the contributions of the exchange and external field for the effective field given by

$$\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) = \ell_{\text{ex}}^2 \, \Delta \boldsymbol{m} + \boldsymbol{h}_{\text{ext}},$$

where $\ell_{\text{ex}} > 0$ denotes the exchange length and $\boldsymbol{h}_{\text{ext}}$ is an externally given field. Throughout this work, we let $\boldsymbol{f}$ denote the external field $\boldsymbol{h}_{\text{ext}}$ for simplicity. To derive a weak formulation, we define the tangent space to $\boldsymbol{P}(\boldsymbol{m})$ by

$$\boldsymbol{T}(\boldsymbol{m}) \coloneqq \{\boldsymbol{\varphi} \in \boldsymbol{\mathcal{H}}^1(\Omega) \ : \ \boldsymbol{m} \cdot \boldsymbol{\varphi} = 0 \text{ a.e.}\} = \{\boldsymbol{\varphi} \in \boldsymbol{\mathcal{H}}^1(\Omega) \ : \ \boldsymbol{P}(\boldsymbol{m})\boldsymbol{\varphi} = \boldsymbol{\varphi}\}.$$

By multiplying (1.6.1) with a test function in the tangent space, integrating over the domain Ω and using the self-adjointness of $\boldsymbol{P}(\boldsymbol{m})$, we obtain

$$\alpha\langle\partial_t\boldsymbol{m},\boldsymbol{\varphi}\rangle+\langle\boldsymbol{m}\times\partial_t\boldsymbol{m},\boldsymbol{\varphi}\rangle=\langle\boldsymbol{P}(\boldsymbol{m})\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}),\boldsymbol{\varphi}\rangle=\langle\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}),\boldsymbol{\varphi}\rangle$$

for all $\boldsymbol{\varphi}\in\boldsymbol{T}(\boldsymbol{m})$. After performing spatial integration by parts on the exchange term, we obtain the following weak formulation: Find $\boldsymbol{m}\in\mathcal{L}^2(0,T;\boldsymbol{V})$ with $\partial_t\boldsymbol{m}\in\mathcal{L}^2(0,T;\boldsymbol{T}(\boldsymbol{m}))$ such that

$$\alpha\langle\partial_t\boldsymbol{m},\boldsymbol{\varphi}\rangle+\langle\boldsymbol{m}\times\partial_t\boldsymbol{m},\boldsymbol{\varphi}\rangle+\ell_{\mathrm{ex}}^2\langle\boldsymbol{\nabla}\boldsymbol{m},\boldsymbol{\nabla}\boldsymbol{\varphi}\rangle=\langle\boldsymbol{f},\boldsymbol{\varphi}\rangle,\tag{1.6.3}$$

for all $\boldsymbol{\varphi}\in\mathcal{L}^2(0,T;\boldsymbol{T}(\boldsymbol{m}))$ for almost every $t\in(0,T)$ with $\boldsymbol{f}\in\mathcal{L}^2(0,T;\boldsymbol{V}')$. In this formulation, the constraint $\partial_t\boldsymbol{m}\cdot\boldsymbol{m}=0$ is implicitly satisfied by the choice of space for the time derivative.

We can derive an alternative weak formulation of equation (1.6.3) by enforcing the normalization constraint using a Lagrange multiplier. More precisely, we can impose the saddle point formulation: Find $\boldsymbol{m}\in\mathcal{L}^2(0,T;\boldsymbol{\mathcal{H}}^1(\Omega))\cap\mathcal{H}^1(0,T;\boldsymbol{\mathcal{L}}^2(\Omega))$ and $\lambda\in\mathcal{L}^2(0,T;\widetilde{\mathcal{H}}^{-1}(\Omega))$, such that

$$\begin{aligned}\alpha\langle\partial_t\boldsymbol{m},\boldsymbol{\varphi}\rangle+\langle\boldsymbol{m}\times\partial_t\boldsymbol{m},\boldsymbol{\varphi}\rangle+\ell_{\mathrm{ex}}^2\langle\boldsymbol{\nabla}\boldsymbol{m},\boldsymbol{\nabla}\boldsymbol{\varphi}\rangle+\langle\boldsymbol{m}\cdot\boldsymbol{\varphi},\lambda\rangle&=\langle\boldsymbol{f},\boldsymbol{\varphi}\rangle,\\ \langle\boldsymbol{m}\cdot\partial_t\boldsymbol{m},\psi\rangle&=0,\end{aligned}\tag{1.6.4}$$

for all $(\boldsymbol{\varphi},\psi)\in\mathcal{L}^2(0,T;\boldsymbol{\mathcal{H}}^1(\Omega))\times\mathcal{L}^2(0,T;\widetilde{\mathcal{H}}^{-1}(\Omega))$ for almost every $t\in(0,T)$.

Let $0=t_0<t_1<\cdots<t_N=T$ be a partition of $[0,T]$ with time-step sizes $\tau_n\coloneqq t_n-t_{n-1}$. We denote by $\boldsymbol{m}^n$ an approximation to $\boldsymbol{m}(t_n)$ and by $\boldsymbol{v}^n$ an approximation to the time derivative $\partial_t\boldsymbol{m}(t_n)$. For time stepping, we employ the linearly implicit backward differentiation formulas (BDF) of order $k\in\{1,\dots,5\}$. We abbreviate the BDF method of order k by BDF(k). The BDF relation for the discrete time derivative is given by

$$\tau_n\boldsymbol{v}^n=\sum_{j=0}^{k}\delta_j\boldsymbol{m}^{n-j},\tag{1.6.5}$$

where the coefficients $\{\delta_j\}_{j=0}^k$ are the usual BDF coefficients (depending on k). For further details, we refer to Section 3.1. Moreover, we introduce an extrapolation $\widehat{\boldsymbol{m}}^n$ based on the known values $\boldsymbol{m}^{n-k},\dots,\boldsymbol{m}^{n-1}$. For instance, we may construct $\widehat{\boldsymbol{m}}^n$ by finding the unique polynomial $\boldsymbol{q}\in\mathbb{P}_{k-1}$ that interpolates the previous values, i.e, $\boldsymbol{q}(t_{n-1-j})=\boldsymbol{m}^{n-1-j}$ for $j=0,\dots,k-1$, and then setting $\widehat{\boldsymbol{m}}^n=\boldsymbol{q}(t_n)/\left|\boldsymbol{q}(t_n)\right|$. For a more extensive discussion on the extrapolation, we refer to Section 4.2.

Furthermore, let $\mathcal{T}_h^n$ be an affinely equivalent, admissible and shape-regular partition of Ω associated with t_n. Let $V_h^n=\mathcal{S}_h^p(\mathcal{T}_h^n)\subset\mathcal{H}^1(\Omega)$ be the space of globally continuous piecewise polynomials of degree $p\geq1$ on $\mathcal{T}_h^n$. Moreover, $\boldsymbol{V}_h^n\coloneqq(V_h^n)^3$. For a function $\boldsymbol{m}\in\boldsymbol{\mathcal{H}}^1(\Omega)$ with

$|\boldsymbol{m}| \neq 0$ a.e., we define the discrete tangent space

$$\boldsymbol{T}_h^n(\boldsymbol{m}) \coloneqq \{\boldsymbol{\varphi}_h^n \in \boldsymbol{V}_h^n \ : \ \langle \boldsymbol{m} \cdot \boldsymbol{\varphi}_h^n, \psi_h^n \rangle = 0 \text{ for all } \psi_h^n \in V_h^n\}.$$

Note that $\boldsymbol{T}_h^n(\boldsymbol{m}) \not\subset \boldsymbol{T}(\boldsymbol{m})$ in general. In contrast to the orthogonal projection $\boldsymbol{P}(\boldsymbol{m})$ on the tangent space $\boldsymbol{T}(\boldsymbol{m})$, we do not have an explicit expression for the orthogonal projection onto the discrete tangent space $\boldsymbol{T}_h^n(\boldsymbol{m})$ [7, Cpt. 5]. Additionally, we would like to point out that the discrete tangent space we use differs from that used in the works [13, 16]. In our approach, we enforce the normalization constraint in an averaged $\mathcal{L}^2$-sense. In contrast, the normalization constraint is satisfied at each node of the finite element mesh in [13, 16]. Then the fully discrete scheme is given by: Find $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ with $\boldsymbol{v}_h^n \in \boldsymbol{T}_h^n(\widehat{\boldsymbol{m}}_h^n)$ such that

$$\alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^n, \boldsymbol{\nabla} \boldsymbol{\varphi}_h^n \rangle = \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n \rangle, \tag{1.6.6}$$

for all $\boldsymbol{\varphi}_h^n \in \boldsymbol{T}_h^n(\widehat{\boldsymbol{m}}_h^n)$, where $\boldsymbol{f}_h^n \coloneqq \boldsymbol{P}_0^n \boldsymbol{f}(t_n)$ with the $\mathcal{L}^2$ projection $\boldsymbol{P}_0^n$ onto $\boldsymbol{V}_h^n$. Utilizing the BDF scheme (1.6.5), we obtain the linear equation for $\boldsymbol{v}_h^n$

$$\alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \ell_{\mathrm{ex}}^2 \frac{\tau_n}{\delta_0} \langle \boldsymbol{\nabla} \boldsymbol{v}_h^n, \boldsymbol{\nabla} \boldsymbol{\varphi}_h^n \rangle = \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n \rangle + \sum_{j=1}^{k} \ell_{\mathrm{ex}}^2 \frac{\delta_j}{\delta_0} \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-j}, \boldsymbol{\nabla} \boldsymbol{\varphi}_h^n \rangle \tag{1.6.7}$$

for all $\boldsymbol{\varphi}_h^n \in \boldsymbol{T}_h^n(\widehat{\boldsymbol{m}}_h^n)$. A possibility to deal with the discrete tangent space in the fully discrete scheme is to introduce again a discrete Lagrange multiplier. Then, we can rewrite (1.6.6) as follows: Find $(\boldsymbol{v}_h^n, \lambda_h^n) \in \boldsymbol{V}_h^n \times V_h^n$ such that

$$\begin{aligned} \alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^n, \boldsymbol{\nabla} \boldsymbol{\varphi}_h^n \rangle + \langle \widehat{\boldsymbol{m}}_h^n \cdot \boldsymbol{\varphi}_h^n, \lambda_h^n \rangle &= \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n \rangle, \\ \langle \widehat{\boldsymbol{m}}_h^n \cdot \boldsymbol{v}_h^n, \psi_h^n \rangle &= 0 \end{aligned} \tag{1.6.8}$$

for all $(\boldsymbol{\varphi}_h^n, \psi_h^n) \in \boldsymbol{V}_h^n \times V_h^n$.

To establish a priori error estimates for a polynomial degree $p \geq 1$ and the BDF(k) scheme with $k \geq 1$, we assume that the solution $\boldsymbol{m}$ and the right-hand side $\boldsymbol{f}$ satisfy the regularity conditions

$$\begin{aligned} \boldsymbol{m} \in \mathcal{C}^{k+1}\big([0,T], \boldsymbol{\mathcal{L}}^\infty(\Omega)\big) \cap \mathcal{C}^1\big([0,T]; \boldsymbol{\mathcal{W}}^{p+1,\infty}(\Omega)\big), \\ \Delta \boldsymbol{m} + \boldsymbol{f} \in \mathcal{C}^0\big([0,T]; \boldsymbol{\mathcal{W}}^{p+1,\infty}(\Omega)\big). \end{aligned} \tag{1.6.9}$$

Higher-order estimates for polynomial degree $p \geq 1$ and BDF order $k \geq 1$ naturally requires additional regularity beyond the weak setting. Consequently, we perform our convergence and stability analysis within the notion of strong solutions as introduced in Definition 3. The existence of such (nontrivial) solutions is guaranteed by Theorem 5 for appropriate initial data.

For the following results, we assume that the family of triangulations $\{\mathcal{T}_h^n\}$ is both regular, quasi-uniform and independent of the time-step n. With the mild CFL-type condition $\tau^k \leq c\sqrt{h}$, the following a priori error estimate establishes the convergence of the fully discrete

scheme (1.6.6) for BDF(k), $k = 1, 2$, as proven in [7, Theorem 3.1].

Theorem 10 (A priori error bound for $k = 1, 2$)**.** *Let $\boldsymbol{m}$ be a strong solution of* (1.5.2) *satisfying the regularity assumptions in* (1.6.9)*. Let $\boldsymbol{m}_h^n$ denote the fully discrete solution of* (1.6.6) *for $k = 1, 2$ and $p \geq 1$. Suppose that τ and h are sufficiently small and satisfy*

$$\tau^k \leq c\sqrt{h}$$

for a sufficiently small constant $c > 0$ independent of h and τ. Then, provided the initial error satisfies a consistent bound, the error is bounded by

$$\|\boldsymbol{m}_h^n - \boldsymbol{m}(t_n)\|_{\boldsymbol{\mathcal{H}}^1} \leq C(\tau^k + h^p),$$

where the constant $C > 0$ is independent of h, τ and n, but depends on α and exponentially on T.

For the higher-order BDF(k) schemes with $k \in \{3, 4, 5\}$, the a priori error analysis requires an additional lower bound on the damping parameter α together with the stronger (but mild) CFL-type condition $\tau^k \leq \overline{C}h$. Under these assumptions, the following a priori error bound is provided in [7, Theorem 3.4].

Theorem 11 (A priori error bound for $k = 3, 4, 5$)**.** *Let $\boldsymbol{m}$ be a strong solution of* (1.5.2) *satisfying the regularity assumptions in* (1.6.9)*. Let $\boldsymbol{m}_h^n$ denote the fully discrete solution of* (1.6.6) *for $k = 3, 4, 5$ and $p \geq 2$. Suppose the damping parameter α is bounded by $\alpha > \alpha_k$, where*

$$\alpha_3 = 0.0913, \qquad \alpha_4 = 0.4041, \qquad \alpha_5 = 4.4348.$$

Further, suppose that τ and h are sufficiently small and satisfy

$$\tau^k \leq \overline{C}h$$

for an arbitrary constant $\overline{C} > 0$. Then, provided the initial error satisfies a consistent bound, the error is bounded by

$$\|\boldsymbol{m}_h^n - \boldsymbol{m}(t_n)\|_{\boldsymbol{\mathcal{H}}^1} \leq C(\tau^k + h^p),$$

where the constant $C > 0$ is independent of h, τ and n, but depends on α and exponentially on $\overline{C}T$.

1.7 Contributions and outline of the dissertation

In this dissertation, we contribute to the efficient numerical simulation of the LLG equation. Specifically, we focus on the reliability of residual-based a posteriori error estimators. Furthermore, we analyze the G-stability of variable step-size BDF(k) formulas from which we

obtain a discrete energy inequality. Finally, we develop a full adaptive algorithm based on the theoretical results.

In Chapter 2 we rigorously derive an upper residual-based a posteriori error bound. We focus on the fully discrete tangent plane scheme introduced in Section 1.6 and employ a Lagrange multiplier to handle the non-convex constraint in a weak $\mathcal{L}^2$-sense. To establish the error estimate, we combine the three-point reconstruction (in time) with the elliptic reconstruction into a time-space reconstruction.

In Chapter 3 we derive a discrete energy inequality for the fully discrete tangent plane scheme introduced in Section 1.6 using the variable time-step size BDF(k) method. This stability property is crucial to ensure that the numerical solution preserves the underlying physical behavior of the LLG equation. We first establish the G-stability of the variable step-size BDF(k) formulas and then utilize the Nevanlinna–Odeh multiplier technique to obtain the desired estimate. The resulting discrete energy inequality is valid under the assumption of a lower bound on the damping coefficient α and an upper bound on the consecutive time-step size ratio. While this analysis is partially based on our previous work [36], we extend the results to BDF(k), $k \in \{1, \dots, 5\}$, provided a single multiplier exists for this method. Finally, we employ discrete orthogonal convolution kernels to transform the discrete energy inequality into an $\mathcal{L}^2$-bound for the discrete solution.

Building upon the theoretical results, we establish in Chapter 4 a space and time adaptive algorithm for the LLG equation. We derive the local truncation error for the variable step-size BDF(k) method and provide corresponding computable approximations. Furthermore, we develop a variable step-size and variable order time-stepping strategy. By combining this approach with residual-based spatial error estimation, we obtain a fully adaptive framework.

In Chapter 5 we verify that the a posteriori error estimators introduced in Chapter 2 converge at the optimal order with respect to either the time-step size or the spatial mesh size. Moreover, we apply the proposed space and time adaptive algorithm to various numerical examples to demonstrate its effectiveness and robustness.

For the convenience of the reader, we provide an extensive list of the abbreviations and symbols at the end of this thesis.

CHAPTER 2

Residual-based a posteriori error estimate

In this chapter, we establish a rigorous upper a posteriori error bound for the fully discrete scheme (2.3.2) based on BDF(2) time-stepping and finite element spaces of arbitrary polynomial degree p, relative to the weak formulation (1.6.3). The effective field considered is

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

Numerical results in Section 5.3 confirm that all error indicators converge at the optimal rate. To the best of the authors knowledge, this is the first rigorous a posteriori error estimate for the LLG equation including the exchange field. In particular, the normalization constraint $\partial_t \boldsymbol{m} \cdot \boldsymbol{m} = 0$ is accounted for in a weak $\mathcal{L}^2$-sense, while higher-order spatial approximations and second-order accuracy in time are admissible.

The main idea is to combine the three-point reconstruction from [6, 96] with the elliptic reconstruction [91, 97] into a time-space reconstruction. A similar concept to the elliptic reconstruction in the literature is the post-processed solution, where an auxiliary approximation is introduced to obtain a posteriori error bounds [51, 52]. Moreover, we follow ideas from [25], where a posteriori estimates for the time-dependent Stokes equations for the BDF(2) scheme are derived utilizing the three-point reconstruction combined with a Stokes reconstruction. Similar to [6, 25], for ease of presentation, we first provide the analysis for uniform time-steps. The extension to variable step sizes is given in Section 2.6.

To date, the mathematical analysis of a posteriori error estimates for the LLG equation has been limited in the literature. A detailed theoretical study by Baňas [23] considers a space-time discretization and a semi-implicit scheme, where the latter is closely related to a method proposed in [106], which is based on

$$\partial_t \boldsymbol{m} - \alpha \Delta \boldsymbol{m} = \alpha |\nabla \boldsymbol{m}|^2 \boldsymbol{m} + \Delta \boldsymbol{m} \times \boldsymbol{m} + \boldsymbol{f}(\boldsymbol{m}), \tag{2.0.1}$$

where $\boldsymbol{f}(\boldsymbol{m}) = \boldsymbol{h}_{\mathrm{ext}} \times \boldsymbol{m} + \alpha(\boldsymbol{h}_{\mathrm{ext}} - (\boldsymbol{m} \cdot \boldsymbol{h}_{\mathrm{ext}})\boldsymbol{m})$. For both of these discretizations, a posteriori error bounds are provided in [23]. Since the constraint $\partial_t \boldsymbol{m} \cdot \boldsymbol{m} = 0$ is not inherited by (2.0.1), the formulation is equivalent to the LLG equation (1.2.10) only under the additional condition $|\boldsymbol{m}| = 1$. Both discretizations in [23], however, do not enforce this constraint. Moreover, although the scheme [23, (3.69)] is first order accurate, it suffers from severe instabilities caused by the violation of the normalization constraint $|\boldsymbol{m}| = 1$ [106, p. 137]. Thus, a penalization is usually required to stabilize such schemes [106]. A posteriori error bounds for a related LLG model with magnetostriction, but without the exchange field, were derived in [21, 24].

2.1 NOTATION

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary, where $d = 2, 3$. We consider the spaces $V \coloneqq \mathcal{H}^1(\Omega)$, $\boldsymbol{V} \coloneqq V^3$, $V' \coloneqq \widetilde{\mathcal{H}}^{-1}(\Omega)$ and the conforming finite element space $V_h^n \coloneqq \mathcal{S}_h^p(\mathcal{T}_h^n) \subset V \subset V'$ for $p \geq 1$ corresponding to the shape regular triangulation $\mathcal{T}_h^n$ of Ω at t_n for $n \geq 0$. Moreover, we consider the tangent spaces according to Section 1.6, namely

$$\boldsymbol{T}(\boldsymbol{m}) = \{\boldsymbol{\varphi} \in \boldsymbol{V} \ : \ \boldsymbol{m} \cdot \boldsymbol{\varphi} = 0 \text{ a.e.}\},$$

as well as the discrete tangent space

$$\boldsymbol{T}_h^n(\boldsymbol{m}) = \{\boldsymbol{\varphi}_h \in \boldsymbol{V}_h^n \ : \ \langle \boldsymbol{m} \cdot \boldsymbol{\varphi}_h, \psi_h \rangle = 0, \quad \forall \psi_h \in V_h^n\}. \tag{2.1.1}$$

Next, we introduce the operator $\boldsymbol{P}(\boldsymbol{m}) \coloneqq \boldsymbol{I} - \boldsymbol{m}\boldsymbol{m}^\top$, as well as the orthogonal projection onto the tangent plane $\boldsymbol{T}(\boldsymbol{m})$ to the unit sphere $\mathbb{S}^2$, defined for $|\boldsymbol{m}| \neq 0$ by

$$\boldsymbol{P}^\perp(\boldsymbol{m}) \coloneqq \boldsymbol{I} - \frac{\boldsymbol{m}\boldsymbol{m}^\top}{|\boldsymbol{m}|^2}.$$

Although $\boldsymbol{P}(\boldsymbol{m}) = \boldsymbol{P}^\perp(\boldsymbol{m})$ if and only if $|\boldsymbol{m}| = 1$, we refer to it as the orthogonal projection onto the tangent plane $\boldsymbol{T}(\boldsymbol{m})$. Later in the analysis, we make use of the fact that $\boldsymbol{P}$ satisfies Lipschitz-type bounds and employ $\boldsymbol{P}$ to measure the deviation of the discrete solution from its tangent plane.

Let $\boldsymbol{P}_0^n$ be the $\mathcal{L}^2$-projection onto $\boldsymbol{V}_h^n$. For a sequence $\{\boldsymbol{m}_h^n\}_{n=-1}^N$ with $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ for $n \geq 0$ and $\boldsymbol{m}_h^{-1} \in \boldsymbol{V}_h^0$, we define the discrete differentials

$$\overline{\partial}\boldsymbol{m}_h^n \coloneqq \overline{\partial}^1 \boldsymbol{m}_h^n \coloneqq \frac{\boldsymbol{m}_h^n - \boldsymbol{m}_h^{n-1}}{\tau} = \frac{\boldsymbol{m}_h^n - \boldsymbol{m}_h^{n-1}}{t_n - t_{n-1}} \tag{2.1.2}$$

for $n \geq 0$ and

$$\overline{\partial}^k \boldsymbol{m}_h^n \coloneqq \overline{\partial}(\overline{\partial}^{k-1} \boldsymbol{m}_h^n) \tag{2.1.3}$$

for $n + 1 \geq k \geq 2$. In general $\overline{\partial}^k \boldsymbol{m}_h^n \notin \boldsymbol{V}_h^n$, we thus define $\overline{\partial}_n^k \boldsymbol{m}_h^n \coloneqq \boldsymbol{P}_0^n \overline{\partial}^k \boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ for $n + 1 \geq k \geq 1$. Further, we define the linear interpolating polynomial

$$\boldsymbol{m}_h(t) = \boldsymbol{m}_h^n + (t - t_n)\overline{\partial}\boldsymbol{m}_h^n \tag{2.1.4}$$

for $t \in I_n \coloneqq (t_{n-1}, t_n]$ and $n \geq 1$, as well as the three-point reconstruction

$$\begin{aligned} \boldsymbol{M}_h(t) &\coloneqq \boldsymbol{m}_h(t) + \frac{1}{2}(t - t_n)(t - t_{n-1})\overline{\partial}^2 \boldsymbol{m}_h^n, \\ \partial_t \boldsymbol{M}_h(t) &= \overline{\partial}^B \boldsymbol{m}_h^n + (t - t_n)\overline{\partial}^2 \boldsymbol{m}_h^n \end{aligned} \tag{2.1.5}$$

for $t \in I_n \coloneqq (t_{n-1}, t_n]$ and $n \geq 2$, where

$$\overline{\partial}^B \boldsymbol{m}_h^n \coloneqq \frac{\tau}{2}\overline{\partial}^2 \boldsymbol{m}_h^n + \overline{\partial}\boldsymbol{m}_h^n, \qquad \overline{\partial}_n^B \boldsymbol{m}_h^n \coloneqq \boldsymbol{P}_0^n \overline{\partial}^B \boldsymbol{m}_h^n \tag{2.1.6}$$

is the BDF(2) time discretization operator. Visual representations of the linear and three-point reconstruction are given in Figure 2.1.1. For $t \in I_1 = (t_0, t_1]$, we define the initial three-point reconstruction

$$\boldsymbol{M}_h(t) = \boldsymbol{m}_h(t) + \frac{1}{2}(t-t_0)(t-t_1)\overline{\partial}^2 \boldsymbol{m}_h^2, \qquad \partial_t \boldsymbol{M}_h(t) = \overline{\partial}\boldsymbol{m}_h^1 + (t-t_{1/2})\overline{\partial}^2 \boldsymbol{m}_h^2, \tag{2.1.7}$$

where $t_{1/2} \coloneqq (t_1 + t_0)/2$. Although the same symbol $\boldsymbol{M}_h$ is used for the three-point reconstruction on I_1, its meaning is evident from the context.

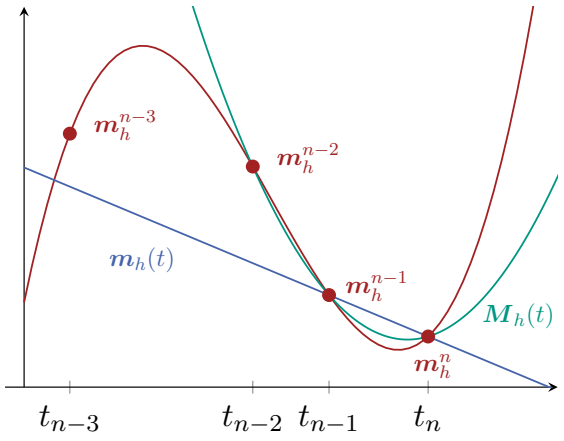


Figure 2.1.1: Visual representation of the linear and the three-point (quadratic) reconstructions $\boldsymbol{m}_h$ (blue) and $\boldsymbol{M}_h$ (green) for given values $\boldsymbol{m}_h^{n-j}$, $j = 0, 1, 2$. The curve of an exemplary solution is given in red. The three-point reconstruction is the unique quadratic polynomial, which interpolates $\boldsymbol{m}_h^{n-j}$ for $j = 0, 1, 2$.

2.2 Preliminaries

We consider the bilinear form $a : \boldsymbol{V} \times \boldsymbol{V} \to \mathbb{R}$ defined by

$$a(\boldsymbol{m}, \boldsymbol{\varphi}) = \langle \boldsymbol{\nabla}\boldsymbol{m}, \boldsymbol{\nabla}\boldsymbol{\varphi} \rangle. \tag{2.2.1}$$

We first introduce the discrete Laplacian for functions in $\boldsymbol{\mathcal{H}}^1(\Omega)$. An analogous definition for functions in $\boldsymbol{H}_0^1(\Omega)$ is given in [26, Def. 1.1].

Definition 12 (Discrete Laplacian)**.** *Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded domain with Lipschitz boundary and $\boldsymbol{v} \in \boldsymbol{\mathcal{H}}^1(\Omega)$. Then the discrete Laplacian $-\Delta_h^n \boldsymbol{v} \in \boldsymbol{V}_h^n \subset \boldsymbol{\mathcal{H}}^1(\Omega)$ is the operator with the property*

$$\langle -\Delta_h^n \boldsymbol{v}, \boldsymbol{\varphi}_h^n \rangle = a(\boldsymbol{v}, \boldsymbol{\varphi}_h^n) - \langle \gamma_N \boldsymbol{v}, \gamma_D \boldsymbol{\varphi}_h^n \rangle, \qquad \forall \boldsymbol{\varphi}_h^n \in \boldsymbol{V}_h^n, \tag{2.2.2}$$

where the trace operators γ_D and γ_N are given in Theorem 1 and Lemma 2.

We proceed with the definition of the elliptic reconstruction associated with the bilinear form a and the finite element space $\boldsymbol{V}_h^n$ [91, 97].

Definition 13 (Elliptic reconstruction)**.** *We define for $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n \subset \boldsymbol{\mathcal{H}}^1(\Omega)$ the elliptic reconstruction $\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n \in \boldsymbol{\mathcal{H}}^1(\Omega)$ by*

$$a(\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n, \boldsymbol{\varphi}) = \langle -\Delta_h^n \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle, \qquad \forall \boldsymbol{\varphi} \in \boldsymbol{\mathcal{H}}^1(\Omega). \tag{2.2.3}$$

A main characteristic of the elliptic reconstruction is that we can use any available a posteriori estimate for the elliptic part of the equation [91]. First, we want to emphasize that the elliptic reconstruction is introduced with homogeneous Neumann boundary conditions. Second, we consider later in Section 2.3 a weak formulation of the LLG equation that likewise enforces homogeneous Neumann boundary conditions. As a consequence, the discrete solution $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ satisfies

$$\langle \gamma_{\mathrm{N}} \boldsymbol{m}_h^n, \gamma_{\mathrm{D}} \boldsymbol{\varphi}_h^n \rangle = 0 \tag{2.2.4}$$

for all $\boldsymbol{\varphi}_h^n \in \boldsymbol{V}_h^n$. For now, let us assume that a function $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ satisfies (2.2.4). Then, the Galerkin orthogonality of $\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n$ follows immediately from (2.2.3) and (2.2.2), namely

$$\langle \boldsymbol{\nabla}(\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n), \boldsymbol{\nabla}\boldsymbol{\varphi}_h^n \rangle = \langle -\Delta_h^n \boldsymbol{m}_h^n, \boldsymbol{\varphi}_h^n \rangle - a(\boldsymbol{m}_h^n, \boldsymbol{\varphi}_h^n) = 0 \tag{2.2.5}$$

for all $\boldsymbol{\varphi}_h^n \in \boldsymbol{V}_h^n$. Thus, partial integration on each triangle, (2.2.3) and subtracting (2.2.5) leads to

$$\begin{aligned}\langle \boldsymbol{\nabla}(\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n), \boldsymbol{\nabla}\boldsymbol{\varphi}\rangle = &\sum_{K\in\mathcal{T}_h^n} \langle (\Delta - \Delta_h^n)\boldsymbol{m}_h^n, \boldsymbol{\varphi} - \boldsymbol{\varphi}_h^n\rangle_K + \sum_{E\in\Sigma_n^\circ} \langle [\![\partial_{\boldsymbol{n}} \boldsymbol{m}_h^n]\!]_E, \boldsymbol{\varphi} - \boldsymbol{\varphi}_h^n\rangle_E \\ &+ \sum_{E\in\Sigma_n\cap\partial\Omega} \langle \partial_{\boldsymbol{n}} \boldsymbol{m}_h^n, \boldsymbol{\varphi} - \boldsymbol{\varphi}_h^n\rangle_E\end{aligned}$$

for all $\boldsymbol{\varphi} \in \boldsymbol{V}$ and $\boldsymbol{\varphi}_h^n \in \boldsymbol{V}_h^n$. Defining the element and edge residuals by

$$R_K(\boldsymbol{m}_h^n) \coloneqq (\Delta - \Delta_h^n)\boldsymbol{m}_h^n$$

for $K \in \mathcal{T}_h^n$ and

$$R_E(\boldsymbol{m}_h^n) \coloneqq \begin{cases} [\![\partial_{\boldsymbol{n}} \boldsymbol{m}_h^n]\!]_E, & \text{for } E \in \Sigma_n^\circ, \\ \partial_{\boldsymbol{n}} \boldsymbol{m}_h^n, & \text{for } E \in \Sigma_n \cap \partial\Omega, \end{cases}$$

we obtain by applying standard a posteriori arguments

$$\|\boldsymbol{\nabla}(\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n)\|^2 \lesssim \sum_{K\in\mathcal{T}_h^n} h_K^2 \|R_K(\boldsymbol{m}_h^n)\|_{\boldsymbol{\mathcal{L}}^2(K)}^2 + \sum_{E\in\Sigma_n} h_E \|R_E(\boldsymbol{m}_h^n)\|_{\boldsymbol{\mathcal{L}}^2(E)}^2.$$

By the Aubin–Nitsche duality method and under suitable regularity assumptions, we can additionally show

$$\|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n\|^2 \lesssim \sum_{K\in\mathcal{T}_h^n} h_K^4 \|R_K(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(K)} + \sum_{E\in\Sigma_n} h_E^3 \|R_E(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(E)}.$$

For details, we refer the interested reader to [4, Theorem 3.1], [118, Prop. 3.8] and [119]. We summarize the above results in the following Lemma.

Lemma 14 (A posteriori bound for the elliptic reconstruction)**.** *For any $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ with $\langle \gamma_N \boldsymbol{m}_h^n, \gamma_D \boldsymbol{\varphi}_h^n\rangle = 0$ for all $\boldsymbol{\varphi}_h^n \in \boldsymbol{V}_h^n$, there holds, under suitable regularity assumptions,*

$$\|(\boldsymbol{I} - \boldsymbol{\mathcal{R}})\boldsymbol{m}_h^n\|^2 \lesssim \eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{L}}^2), \qquad \|(\boldsymbol{I} - \boldsymbol{\mathcal{R}})\boldsymbol{m}_h^n\|^2_{\boldsymbol{\mathcal{H}}^1} \lesssim \eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{H}}^1), \tag{2.2.6}$$

where η is defined by

$$\begin{aligned} \eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{L}}^2) &:= \sum_{K\in\mathcal{T}_h^n} h_K^4 \|R_K(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(K)} + \sum_{E\in\Sigma_n} h_E^3 \|R_E(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(E)}, \\ \eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{H}}^1) &:= \sum_{K\in\mathcal{T}_h^n} h_K^2 \|R_K(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(K)} + \sum_{E\in\Sigma_n} h_E \|R_E(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(E)}. \end{aligned} \tag{2.2.7}$$

We also require a posteriori bounds in the $\boldsymbol{\mathcal{L}}^\infty$-norm for the elliptic reconstruction. Again, since the elliptic reconstruction satisfies an associated elliptic problem, any existing a posteriori $\boldsymbol{\mathcal{L}}^\infty$-estimate transfers directly to the reconstruction. For instance, Nochetto et al. derived a posteriori error estimates in the $\mathcal{L}^\infty$-norm for monotone semi-linear elliptic problems on a bounded, polyhedral domain in [101, Theorem 4.2]. Similarly, in [54, Theorem 3.1], an a posteriori bound in the $\mathcal{L}^\infty$-norm of the gradient of the error of piecewise linear finite element approximations is established on convex polyhedral domains for quasilinear elliptic problems.

Assumption 15. *Let $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ and $\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n \in \boldsymbol{V}$ be the elliptic reconstruction defined in* (2.2.3)*. Then, we assume that there exist a posteriori error estimators $\eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{L}}^\infty)$ and $\eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{W}}^{1,\infty})$, such that*

$$\|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n\|^2_{\boldsymbol{\mathcal{L}}^\infty} \lesssim \eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{L}}^\infty), \qquad \|\boldsymbol{\nabla}(\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^\infty} \lesssim \eta(\boldsymbol{m}_h^n; \boldsymbol{\mathcal{W}}^{1,\infty}). \tag{2.2.8}$$

Note that we may have to impose additional assumptions on the domain Ω to ensure the validity of the a posteriori estimates (2.2.6) and (2.2.8), e.g., smoothness of the boundary or convexity.

Further, we require a Lipschitz-type bound for the projection onto the tangent plane. The following Lemma is an extension of [7, Lemma 4.1] to the case $|\boldsymbol{m}| \neq 1$.

Lemma 16 (Lipschitz-type bound)**.** *Let $\widetilde{\boldsymbol{m}}, \boldsymbol{v} \in \boldsymbol{\mathcal{H}}^1(\Omega) \cap \boldsymbol{\mathcal{W}}^{1,\infty}(\Omega)$ and $\boldsymbol{m} \in \boldsymbol{\mathcal{H}}^1(\Omega) \cap \boldsymbol{\mathcal{L}}^\infty(\Omega)$. Then, the projection onto the tangent plane $\boldsymbol{P}(\boldsymbol{m}) = \boldsymbol{I} - \boldsymbol{m}\boldsymbol{m}^\top$ satisfies*

$$\begin{aligned} \|(\boldsymbol{P}(\widetilde{\boldsymbol{m}}) - \boldsymbol{P}(\boldsymbol{m}))\boldsymbol{v}\| &\leq (\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty})\|\boldsymbol{v}\|_{\boldsymbol{\mathcal{L}}^\infty}\|\widetilde{\boldsymbol{m}} - \boldsymbol{m}\|, \\ \|\boldsymbol{\nabla}(\boldsymbol{P}(\widetilde{\boldsymbol{m}}) - \boldsymbol{P}(\boldsymbol{m}))\boldsymbol{v}\| &\leq (2\|\boldsymbol{v}\|_{\boldsymbol{\mathcal{L}}^\infty}\|\boldsymbol{\nabla}\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\boldsymbol{\nabla}\boldsymbol{v}\|_{\boldsymbol{\mathcal{L}}^\infty}(\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty}))\|\widetilde{\boldsymbol{m}} - \boldsymbol{m}\| \end{aligned}$$

$$
\begin{aligned}
&\quad + 2\|\boldsymbol{v}\|_{\boldsymbol{\mathcal{L}}^\infty}(\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty} + 2\|\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty})\|\boldsymbol{\nabla}(\widetilde{\boldsymbol{m}} - \boldsymbol{m})\| \\
&\leq \gamma(\boldsymbol{m}, \widetilde{\boldsymbol{m}}, \boldsymbol{v})\|\widetilde{\boldsymbol{m}} - \boldsymbol{m}\|_{\boldsymbol{\mathcal{H}}^1},
\end{aligned}
$$

where γ is defined by

$$
\gamma(\boldsymbol{m}, \widetilde{\boldsymbol{m}}, \boldsymbol{v}) \coloneqq 2\|\boldsymbol{v}\|_{\boldsymbol{\mathcal{L}}^\infty}(\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty} + 2\|\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\boldsymbol{\nabla}\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty}) + \|\boldsymbol{\nabla}\boldsymbol{v}\|_{\boldsymbol{\mathcal{L}}^\infty}(\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty}).
$$

If $\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty} = 1$, then $\gamma(\boldsymbol{m}, \widetilde{\boldsymbol{m}}, \boldsymbol{v}) = \gamma(\widetilde{\boldsymbol{m}}, \boldsymbol{v})$.

Proof. Setting $\boldsymbol{e} = \boldsymbol{m} - \widetilde{\boldsymbol{m}}$, we rewrite

$$
(\boldsymbol{P}(\widetilde{\boldsymbol{m}}) - \boldsymbol{P}(\boldsymbol{m}))\boldsymbol{v} = -(\boldsymbol{m}\boldsymbol{e}^\top + \boldsymbol{e}\widetilde{\boldsymbol{m}})\boldsymbol{v}
$$

and, using the product and $\boldsymbol{m} = \boldsymbol{e} + \widetilde{\boldsymbol{m}}$,

$$
\begin{aligned}
\partial_i(\boldsymbol{P}(\widetilde{\boldsymbol{m}}) - \boldsymbol{P}(\boldsymbol{m}))\boldsymbol{v} = &-(\partial_i\boldsymbol{e}\boldsymbol{e}^\top + \boldsymbol{e}\partial_i\boldsymbol{e}^\top + \partial_i\widetilde{\boldsymbol{m}}\boldsymbol{e}^\top + \widetilde{\boldsymbol{m}}\partial_i\boldsymbol{e}^\top + \partial_i\boldsymbol{e}\widetilde{\boldsymbol{m}}^\top + \boldsymbol{e}\partial_i\widetilde{\boldsymbol{m}}^\top)\boldsymbol{v} \\
&+ (\boldsymbol{m}\boldsymbol{e}^\top + \boldsymbol{e}\widetilde{\boldsymbol{m}}^\top)\partial_i\boldsymbol{v}
\end{aligned}
$$

and $\|\boldsymbol{e}\|_{\boldsymbol{\mathcal{L}}^\infty} \leq \|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty}$ yields the bounds. □

2.3 Model equation

In this chapter, we consider the tangent plane scheme from [7] for BDF(2) and polynomial degree $p \geq 1$, introduced in Section 1.6. Since the notation is already quite involved, we assume $\ell^2_{\mathrm{ex}} = 1$. However, the proof can be generalized directly to arbitrary $\ell^2_{\mathrm{ex}} > 0$. For the convenience of the reader, we start by recalling the weak formulation (1.6.3) and the corresponding discrete equation (1.6.6) rewritten with the bilinear form a introduced in (2.2.1). We consider the model equation: Find $\boldsymbol{m} \in \mathcal{L}^2(0,T;\boldsymbol{V})$ with $\partial_t\boldsymbol{m} \in \mathcal{L}^2(0,T;\boldsymbol{T}(\boldsymbol{m}))$ such that

$$
\alpha\langle\partial_t\boldsymbol{m}, \boldsymbol{\varphi}\rangle + \langle\boldsymbol{m} \times \partial_t\boldsymbol{m}, \boldsymbol{\varphi}\rangle + a(\boldsymbol{m}, \boldsymbol{\varphi}) = \langle\boldsymbol{f}, \boldsymbol{\varphi}\rangle, \qquad \forall\boldsymbol{\varphi} \in \mathcal{L}^2(0,T;\boldsymbol{T}(\boldsymbol{m})) \tag{2.3.1}
$$

with $\boldsymbol{f} \in \mathcal{L}^2(0,T;\boldsymbol{V}')$ and natural boundary conditions $\partial_{\boldsymbol{n}}\boldsymbol{m} = \boldsymbol{0}$. Then, the fully discrete scheme is given by: Find $\boldsymbol{m}_h^n$ with $\boldsymbol{v}_h^n \coloneqq \overline{\partial}^B\boldsymbol{m}_h^n \in \boldsymbol{T}_h^n(\widehat{\boldsymbol{m}}_h^n)$ such that

$$
\alpha\langle\boldsymbol{v}_h^n, \boldsymbol{\varphi}_h\rangle + \langle\widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h\rangle + a(\boldsymbol{m}_h^n, \boldsymbol{\varphi}_h) = \langle\boldsymbol{f}_h^n, \boldsymbol{\varphi}_h\rangle, \qquad \forall\boldsymbol{\varphi}_h \in \boldsymbol{T}_h^n(\widehat{\boldsymbol{m}}_h^n), \tag{2.3.2}
$$

where $\boldsymbol{f}_h^n \coloneqq \boldsymbol{P}_0^n\boldsymbol{f}(t_n)$ and $\widehat{\boldsymbol{m}}_h^n$ is a predictor for $\boldsymbol{m}_h^n$. Note that our analysis does not rely on any specific structure of $\widehat{\boldsymbol{m}}_h^n$ and does only assume that $\widehat{\boldsymbol{m}}_h^n \in \boldsymbol{V}$, i.e., $\widehat{\boldsymbol{m}}_h^n \notin \boldsymbol{V}_h^n$ is admissible. In general, we first compute an extrapolation $\boldsymbol{q}_h^n$ for $\boldsymbol{m}_h^n$ from a composition of previously computed values $\boldsymbol{m}_h^{n-1}, \dots, \boldsymbol{m}_h^{n-j}$, $j \geq 1$, cf. Section 1.6. Preferably, we then normalize the approximation $\boldsymbol{q}_h^n$ to employ the predictor $\widehat{\boldsymbol{m}}_h^n = \boldsymbol{q}_h^n/|\boldsymbol{q}_h^n|$ as in [7, (2.1)]. However, this normalization leads to $\widehat{\boldsymbol{m}}_h^n \notin \boldsymbol{V}_h^n$. In practice, we may only normalize the predictor at nodal values and thus ensure that the predictor remains in $\boldsymbol{V}_h^n$.

The equations (2.3.1) and (2.3.2) are equivalent to the saddle point problem: Find $(\boldsymbol{m}, \lambda) \in (\mathcal{L}^2(0,T;\boldsymbol{V}) \cap \mathcal{H}^1(0,T;\boldsymbol{\mathcal{L}}^2(\Omega))) \times \mathcal{L}^2(0,T;V')$

$$\begin{aligned} \alpha\langle \partial_t \boldsymbol{m}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m} \times \partial_t \boldsymbol{m}, \boldsymbol{\varphi}\rangle + a(\boldsymbol{m}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\lambda, \boldsymbol{\varphi}) &= \langle \boldsymbol{f}, \boldsymbol{\varphi}\rangle, \\ b^{\boldsymbol{m}}(\psi, \partial_t \boldsymbol{m}) &= 0 \end{aligned} \tag{2.3.3}$$

for all $(\boldsymbol{\varphi}, \psi) \in \mathcal{L}^2(0,T;\boldsymbol{V}) \times \mathcal{L}^2(0,T;V')$ and find $(\boldsymbol{m}_h^n, \lambda_h^n) \in \boldsymbol{V}_h^n \times V_h^n \subset \boldsymbol{V} \times V'$ such that

$$\begin{aligned} \alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h\rangle + a(\boldsymbol{m}_h^n, \boldsymbol{\varphi}_h^n) + b^{\widehat{\boldsymbol{m}}_h^n}(\lambda_h^n, \boldsymbol{\varphi}_h^n) &= \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n\rangle, \\ b^{\widehat{\boldsymbol{m}}_h^n}(\psi_h^n, \boldsymbol{v}_h^n) &= 0 \end{aligned} \tag{2.3.4}$$

for all $(\boldsymbol{\varphi}_h^n, \psi_h^n) \in \boldsymbol{V}_h^n \times V_h^n$, where $b^{\boldsymbol{m}}$ is given by

$$b^{\boldsymbol{m}}(\lambda, \boldsymbol{\varphi}) \coloneqq \langle \lambda, \boldsymbol{m} \cdot \boldsymbol{\varphi}\rangle. \tag{2.3.5}$$

For the initialization of the numerical scheme we require the given initial condition $\boldsymbol{m}^0 \in \boldsymbol{\mathcal{H}}^1(\Omega)$ with $|\boldsymbol{m}^0| = 1$ a.e. in Ω. Then, we set $\boldsymbol{m}_h^0 \coloneqq \boldsymbol{P}_0^0 \boldsymbol{m}^0$. Further, we require λ_h^0 to compute the first time step using the trapezoidal scheme (2.3.7). We determine $(\boldsymbol{v}_h^0, \lambda_h^0) \in \boldsymbol{V}_h^0 \times V_h^0$ by computing

$$\begin{aligned} \alpha\langle \boldsymbol{v}_h^0, \boldsymbol{\varphi}_h^0\rangle + \langle \boldsymbol{m}_h^0 \times \boldsymbol{v}_h^0, \boldsymbol{\varphi}_h^0\rangle + b^{\boldsymbol{m}_h^0}(\lambda_h^0, \boldsymbol{\varphi}_h^0) &= \langle \boldsymbol{f}_h^0, \boldsymbol{\varphi}_h^0\rangle - a(\boldsymbol{m}_h^0, \boldsymbol{\varphi}_h^0), \\ b^{\boldsymbol{m}_h^0}(\psi_h^0, \boldsymbol{v}_h^0) &= 0, \end{aligned} \tag{2.3.6}$$

for all $(\boldsymbol{\varphi}_h^0, \psi_h^0) \in \boldsymbol{V}_h^0 \times V_h^0$. Setting $\widehat{\boldsymbol{m}}_h^0 = \boldsymbol{m}_h^0$, we treat the first time step using the trapezoidal rule as follows: Find $(\boldsymbol{m}_h^1, \lambda_h^1) \in \boldsymbol{V}_h^1 \times V_h^1$ such that

$$\begin{aligned} \alpha\langle \overline{\partial} \boldsymbol{m}_h^1, \boldsymbol{\varphi}_h^1\rangle + \langle \widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial} \boldsymbol{m}_h^1, \boldsymbol{\varphi}_h^1\rangle + a(\boldsymbol{m}_h^{1/2}, \boldsymbol{\varphi}_h^1) + b^{\widehat{\boldsymbol{m}}_h^{1/2}}(\lambda_h^{1/2}, \boldsymbol{\varphi}_h^1) &= \langle \boldsymbol{f}_h^{1/2}, \boldsymbol{\varphi}_h^1\rangle, \\ b^{\widehat{\boldsymbol{m}}_h^{1/2}}(\psi_h^1, \overline{\partial} \boldsymbol{m}_h^1) &= 0, \end{aligned} \tag{2.3.7}$$

for all $(\boldsymbol{\varphi}_h^1, \psi_h^1) \in \boldsymbol{V}_h^1 \times V_h^1$, where the superscript $1/2$ denotes the average of the values at t_0 and t_1, e.g., $\boldsymbol{f}_h^{1/2} = \frac{1}{2}(\boldsymbol{f}_h^1 + \boldsymbol{f}_h^0)$. Possible choices for the predictor $\widehat{\boldsymbol{m}}_h^1$ are, for instance, the normalized first-order approximation $\widehat{\boldsymbol{m}}_h^1 = \widehat{\boldsymbol{m}}_h^0 / |\widehat{\boldsymbol{m}}_h^0|$ or the normalized second-order approximation $\widehat{\boldsymbol{m}}_h^1 = (\widehat{\boldsymbol{m}}_h^0 + \tau \boldsymbol{v}_h^0)/|\widehat{\boldsymbol{m}}_h^0 + \tau \boldsymbol{v}_h^0|$.

Given enough regularity of the weak solution $\boldsymbol{m}$ for the saddle point formulation (2.3.3), the following inf-sup condition holds.

Lemma 17 (Inf-sup condition)**.** *Let $\boldsymbol{m} \in \boldsymbol{\mathcal{W}}^{1,\infty}(\Omega)$ with $|\boldsymbol{m}| = 1$ almost everywhere. Then, for $b^{\boldsymbol{m}} : V' \times \boldsymbol{V} \to \mathbb{R}$ defined in* (2.3.5)*, the inf-sup condition*

$$\beta(\boldsymbol{m}) \|\lambda\|_{V'} \leq \sup_{\boldsymbol{\varphi} \in \boldsymbol{V} \setminus \{\boldsymbol{0}\}} \frac{b^{\boldsymbol{m}}(\lambda, \boldsymbol{\varphi})}{\|\boldsymbol{\varphi}\|_{\boldsymbol{V}}} = \sup_{\boldsymbol{\varphi} \in \boldsymbol{V} \setminus \{\boldsymbol{0}\}} \frac{\langle \lambda, \boldsymbol{m} \cdot \boldsymbol{\varphi}\rangle}{\|\boldsymbol{\varphi}\|_{\boldsymbol{V}}}$$

holds, where β is defined by

$$\beta(\boldsymbol{m}) \coloneqq \frac{1}{\sqrt{3}\max\{1, \|\boldsymbol{\nabla}\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty}\}}.$$

Proof. Using $\boldsymbol{m}\cdot\boldsymbol{m} = 1$ a.e. leads to

$$\|\lambda\|_{V'} = \sup_{\zeta\in V\setminus\{0\}} \frac{\langle\lambda,\zeta\rangle}{\|\zeta\|_V} = \sup_{\zeta\in V\setminus\{0\}} \frac{\langle\lambda,\boldsymbol{m}\cdot\zeta\boldsymbol{m}\rangle}{\|\zeta\|_V}.$$

Since $\boldsymbol{m}\in\boldsymbol{\mathcal{W}}^{1,\infty}(\Omega)$ and $|\boldsymbol{m}| = 1$ a.e., we estimate with Young's inequality

$$\begin{aligned}\|\zeta\boldsymbol{m}\|_{\boldsymbol{V}}^2 &= \int_\Omega \left(|\zeta\boldsymbol{m}|^2 + |\boldsymbol{\nabla}(\zeta\boldsymbol{m})|^2\right)\mathrm{d}\boldsymbol{x}\\ &\le \int_\Omega \left(|\zeta|^2|\boldsymbol{m}|^2 + 2|\zeta|^2|\boldsymbol{\nabla}\boldsymbol{m}|^2 + 2|\boldsymbol{\nabla}\zeta|^2|\boldsymbol{m}|^2\right)\mathrm{d}\boldsymbol{x}\\ &\le 3\max\{1, \|\boldsymbol{\nabla}\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty}^2\}\|\zeta\|_V^2 = \frac{1}{\beta(\boldsymbol{m})^2}\|\zeta\|_V^2.\end{aligned} \tag{2.3.8}$$

Finally, applying (2.3.8) we conclude

$$\|\lambda\|_{V'} = \sup_{\zeta\in V\setminus\{0\}} \frac{\langle\lambda,\boldsymbol{m}\cdot\zeta\boldsymbol{m}\rangle}{\|\zeta\|_V} \le \frac{1}{\beta(\boldsymbol{m})}\sup_{\zeta\in V\setminus\{0\}} \frac{\langle\lambda,\boldsymbol{m}\cdot\zeta\boldsymbol{m}\rangle}{\|\zeta\boldsymbol{m}\|_V} \le \frac{1}{\beta(\boldsymbol{m})}\sup_{\boldsymbol{\varphi}\in\boldsymbol{V}\setminus\{\boldsymbol{0}\}} \frac{\langle\lambda,\boldsymbol{m}\cdot\boldsymbol{\varphi}\rangle}{\|\boldsymbol{\varphi}\|_{\boldsymbol{V}}}.$$

□

On the other hand, discrete inf-sup conditions for $\boldsymbol{m}\in\boldsymbol{\mathcal{W}}^{1,\infty}(\Omega)$ with $|\boldsymbol{m}| = 1$ almost everywhere on a quasi-uniform finite element triangulation are provided in [7, p. 1013]. First, the discrete inf-sup condition

$$\inf_{q_h\in V_h^n}\sup_{\boldsymbol{v}_h\in\boldsymbol{V}_h^n} \frac{\langle q_h,\boldsymbol{m}\cdot\boldsymbol{v}_h\rangle}{\|\boldsymbol{v}_h\|_{\boldsymbol{\mathcal{H}}^s}\|q_h\|_{\mathcal{H}^{-s}}} > 0$$

holds uniformly in h for $s\in\{-1,0,1\}$. Second, for the discrete tangent space (2.1.1), the inf-sup condition

$$\inf_{\boldsymbol{w}_h\in\boldsymbol{T}_h^n(\boldsymbol{m})}\sup_{\boldsymbol{v}_h\in\boldsymbol{T}_h^n(\boldsymbol{m})} \frac{\langle\boldsymbol{v}_h,\boldsymbol{w}_h\rangle}{\|\boldsymbol{v}_h\|_{\boldsymbol{\mathcal{H}}^s}\|\boldsymbol{w}_h\|_{\boldsymbol{\mathcal{H}}^{-s}}} > 0$$

holds uniformly in h for $s\in\{-1,0,1\}$.

2.4 A POSTERIORI ERROR ESTIMATE

Before presenting the main result of this chapter, we first introduce the a posteriori error estimators that we obtain later in our analysis.

Definition 18 (A posteriori error estimators)**.** *Let* $\{\boldsymbol{m}_h^n\}_{n=0}^N$, $\{\widehat{\boldsymbol{m}}_h^n\}_{n=0}^N$, $\{\boldsymbol{f}_h^n\}_{n=0}^N$, $\{\lambda_h^n\}_{n=0}^N$ *be sequences with* $\boldsymbol{m}_h^n, \boldsymbol{f}_h^n \in \boldsymbol{V}_h^n$, $\lambda_h^n \in V_h^n$ *and* $\widehat{\boldsymbol{m}}_h^n \in V$ *for all* $n \geq 0$ *and define*

$$\alpha \boldsymbol{m}_h^{-1} \coloneqq \alpha \boldsymbol{m}_h^1 - 2\tau(\boldsymbol{f}_h^0 + \Delta_h^0 \boldsymbol{m}_h^0 - \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^0 \times \overline{\partial} \boldsymbol{m}_h^1) - \boldsymbol{P}_0^1(\lambda_h^0 \widehat{\boldsymbol{m}}_h^0)). \tag{2.4.1}$$

Then, we introduce the ***time error estimators***

$$\begin{aligned} \mathcal{E}_1 &\coloneqq \max_{2\leq n\leq N} \|\boldsymbol{\nabla} \overline{\partial}_n^2 \boldsymbol{m}_h^n\|^2, & \mathcal{E}_2 &\coloneqq \sum_{n=2}^N \frac{1}{\tau} \|\overline{\partial}_n^2 \boldsymbol{m}_h^n - \overline{\partial}_{n-1}^2 \boldsymbol{m}_h^{n-1}\|^2, \\ \mathcal{E}_3 &\coloneqq \sum_{n=1}^N \tau(\|\overline{\partial} \boldsymbol{m}_h^n \overline{\partial} \lambda_h^n\|^2 + \|\overline{\partial} \widehat{\boldsymbol{m}}_h^n \overline{\partial} \lambda_h^n\|^2), & \mathcal{E}_6 &\coloneqq \sum_{n=2}^N \tau \|\overline{\partial}^2 \boldsymbol{m}_h^n\|^2, \\ \mathcal{E}_4 &\coloneqq \sum_{n=2}^N \tau\Big(\|\overline{\partial} \boldsymbol{m}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n\|^2 + \|\overline{\partial} \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n\|^2 + \|\widehat{\boldsymbol{m}}_h^{n-1} \times \overline{\partial}^3 \boldsymbol{m}_h^n\|^2\Big), \end{aligned} \tag{2.4.2}$$

the ***reconstruction error estimator***

$$\mathcal{E}_5 \coloneqq \sum_{n=2}^N \tau \|\overline{\partial}^2(-\Delta_h^n) \boldsymbol{m}_h^n\|^2,$$

the ***projection error estimator***

$$\mathcal{P} \coloneqq \int_0^{t_N} \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{M}_h)) \partial_t \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{H}}^1}^2 \,\mathrm{d}t,$$

where $\boldsymbol{M}_h$ *is defined in* (2.1.5) - (2.1.7)*, the* ***space error estimators***

$$\Lambda_1 \coloneqq \max_{0\leq n\leq N} \eta(\boldsymbol{m}_h^n; \mathcal{H}^1), \qquad \Lambda_2 \coloneqq \sum_{n=1}^N \tau\, \eta(\overline{\partial}_n \boldsymbol{m}_h^n; \mathcal{H}^1),$$
$$\Lambda_3 \coloneqq \sum_{n=0}^N \tau\, \eta(\boldsymbol{m}_h^n; \mathcal{L}^2),$$

where η *is the elliptic a posteriori estimator defined in* (2.2.7)*, the* ***data approximation error estimators***

$$\mathcal{F}_1 \coloneqq \int_{I_1} \|\boldsymbol{f}_h - \boldsymbol{f}\|^2 \,\mathrm{d}t, \qquad \mathcal{F}_2 \coloneqq \sum_{n=2}^N \int_{I_n} \|\boldsymbol{f}_h - \boldsymbol{f}\|^2 \,\mathrm{d}t,$$

the ***changing mesh estimators***

$$\Xi_1 \coloneqq \sum_{n=2}^N \tau \|(\boldsymbol{I} - \boldsymbol{\mathcal{R}})(\overline{\partial} - \overline{\partial}_n) \boldsymbol{m}_h^n\|^2, \qquad \Xi_2 \coloneqq \tau^4 \max_{2\leq n\leq N} \|\boldsymbol{\nabla}(\overline{\partial}^2 - \overline{\partial}_n^2) \boldsymbol{m}_h^n\|^2,$$

$$\Xi_3 \coloneqq \sum_{n=2}^{N} \Big(\tau \|(\overline{\partial}^B - \overline{\partial}_n^B) \boldsymbol{m}_h^n\|^2 + \tau^3 \|\overline{\partial}((\overline{\partial} - \overline{\partial}_n)\boldsymbol{m}_h^n)\|^2 \Big),$$

the ***finite element space conforming estimators***

$$\begin{aligned}
\mathcal{C}_1 &\coloneqq \tau \|(\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2} \lambda_h^{1/2})\|^2 + \tau \|(\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial} \boldsymbol{m}_h^1)\|^2, \\
\mathcal{C}_2 &\coloneqq \sum_{n=2}^{N} \Big(\tau \|(\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n)\|^2 + \tau \|(\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n)\|^2 \\
&\qquad + \tau^3 \|\overline{\partial}((\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n))\|^2 + \tau^3 \|\overline{\partial}((\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n))\|^2 \Big),
\end{aligned}$$

the ***extrapolation error estimators***

$$\begin{aligned}
\mathcal{Q}_1 &\coloneqq \int_{I_1} (\|(\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \times \partial_t \boldsymbol{M}_h\|^2 + \|(\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^1\|^2) \, \mathrm{d}t + \tau^3 \|(\overline{\partial} \boldsymbol{m}_h^1 - \overline{\partial} \widehat{\boldsymbol{m}}_h^1) \times \overline{\partial} \boldsymbol{m}_h^1\|^2 \\
&\qquad + \tau^2 \|(\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1)\overline{\partial} \lambda_h^1\|^2 \\
\mathcal{Q}_2 &\coloneqq \sum_{n=2}^{N} \Big(\int_{I_n} (\|(\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n) \times \partial_t \boldsymbol{M}_h\|^2 + \|(\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^n\|^2) \, \mathrm{d}t \\
&\qquad + \tau^3 \|(\overline{\partial} \boldsymbol{m}_h^n - \overline{\partial} \widehat{\boldsymbol{m}}_h^n) \times \overline{\partial}^B \boldsymbol{m}_h^n\|^2 + \tau^3 \|(\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n)\overline{\partial} \lambda_h^n\|^2 \Big),
\end{aligned}$$

and the ***initial error estimators***

$$\begin{aligned}
\mathcal{I}_1 &\coloneqq \tau^3 \|\boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^1)\|^2 + \tau^5 \|\boldsymbol{P}_0^1(\overline{\partial} \lambda_h^1 \overline{\partial} \widehat{\boldsymbol{m}}_h^1)\|^2, \qquad \mathcal{I}_2 \coloneqq \tau^3 \|\boldsymbol{\Psi}_0\|^2, \\
\mathcal{I}_3 &\coloneqq \tau^5 \|\overline{\partial} \boldsymbol{m}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^2\|^2 + \tau^5 \|\overline{\partial} \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1\|^2 + \tau^5 \|\overline{\partial} \boldsymbol{m}_h^1 \overline{\partial} \lambda_h^1\|^2,
\end{aligned} \tag{2.4.3}$$

where $\boldsymbol{\Psi}_0 \coloneqq \alpha \overline{\partial}^2 \boldsymbol{m}_h^2 + \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \times \overline{\partial} \boldsymbol{m}_h^1 + \widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^2 - \overline{\partial}(\Delta_h^1 \boldsymbol{m}_h^1) + \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1 + \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \lambda_h^1 - \overline{\partial} \boldsymbol{f}_h^1$.

The main result is given by the following theorem.

Theorem 19 (A posteriori error estimate)**.** *Let* $\boldsymbol{m}$ *denote the exact solution of* (2.3.1)*, and let* $\boldsymbol{m}_h^n$ *be the solution of* (2.3.2) *for* $n = 2, \dots, N$*, with* $\boldsymbol{m}_h^1$ *given by* (2.3.7)*. Then*

$$\begin{aligned}
\|\boldsymbol{\nabla}(\boldsymbol{m} - \boldsymbol{m}_h)\|^2_{\mathcal{L}^\infty(t_0, t_N; \boldsymbol{\mathcal{L}}^2(\Omega))} \lesssim \mathcal{F}_1 + \mathcal{F}_2 + \tau^4(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4 + \mathcal{E}_5 + \mathcal{E}_6) + \Lambda_1 + \Lambda_2 + \Lambda_3 \\
+ \mathcal{P} + \Xi_1 + \Xi_2 + \Xi_3 + \mathcal{C}_1 + \mathcal{C}_2 + \mathcal{Q}_1 + \mathcal{Q}_2 + \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3,
\end{aligned} \tag{2.4.4}$$

where the estimators are defined in Definition 18. The hidden constant depends only on α*,* $\|\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}$*,* $\|\partial_t \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}$*,* $\eta(\boldsymbol{m}_h^n; \mathcal{W}^{1,\infty})$*, and* $\|\lambda_h\|_{\mathcal{L}^\infty}$*.*

Before presenting the proof of Theorem 19, we first introduce appropriate time-space reconstructions and derive a corresponding parabolic error equation. Based on the elliptic reconstruction, we define the linear-in-time reconstruction

$$\boldsymbol{w}(t) = \boldsymbol{\mathcal{R}} \boldsymbol{m}_h^n + (t - t_n) \overline{\partial} \boldsymbol{\mathcal{R}} \boldsymbol{m}_h^n, \tag{2.4.5}$$

for $t \in I_n$, $n \geq 1$, and the quadratic (three-point) time-space reconstruction

$$\boldsymbol{W}(t) = \boldsymbol{w}(t) + \tfrac{1}{2}(t - t_n)(t - t_{n-1})\overline{\partial}^2 \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n, \tag{2.4.6}$$

for $t \in I_n$, $n \geq 2$. For the initial interval I_1, we set

$$\boldsymbol{W}(t) = \boldsymbol{w}(t) + \tfrac{1}{2}(t - t_0)(t - t_1)\overline{\partial}^2 \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^2.$$

For the Lagrangian multiplier, we employ a linear-in-time reconstruction, namely

$$\lambda_h(t) = \lambda_h^n + (t - t_n)\overline{\partial}\lambda_h^n, \tag{2.4.7}$$

for $t \in I_n$, $n \geq 1$.

As a first step, we derive the error equation using both the three-point reconstruction $\boldsymbol{M}_h$ and the time-space reconstruction $\boldsymbol{w}$. To this end, we introduce the error variables

$$\boldsymbol{\mathcal{E}} \coloneqq \boldsymbol{M}_h - \boldsymbol{m}, \qquad \boldsymbol{\epsilon} \coloneqq \boldsymbol{m}_h - \boldsymbol{m}, \qquad \widetilde{\boldsymbol{\mathcal{E}}} \coloneqq \boldsymbol{W} - \boldsymbol{m}, \qquad \widetilde{\boldsymbol{\epsilon}} \coloneqq \boldsymbol{w} - \boldsymbol{m}, \qquad \varepsilon \coloneqq \lambda_h - \lambda. \tag{2.4.8}$$

In the parabolic error equation below, the term $\langle \boldsymbol{\epsilon} \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle$ does not involve a time derivative of $\boldsymbol{m}$, thus a linear-in-time reconstruction is sufficient. Employing the three-point reconstruction in this term merely introduces additional a posteriori terms without providing any advantage.

Lemma 20 (Parabolic error equation)**.** *Let* $(\boldsymbol{m}, \lambda)$ *be the exact solution of* (2.3.3) *and* $(\boldsymbol{m}_h^n, \lambda_h^n)$ *the solution of* (2.3.4)*. Then, for* $t \in I_n$ *and* $n \geq 2$*, we have*

$$\alpha\langle \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi} \rangle + \langle \boldsymbol{m} \times \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi} \rangle + \langle \boldsymbol{\epsilon} \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi}) = \langle \boldsymbol{r}_h, \boldsymbol{\varphi} \rangle, \tag{2.4.9}$$

for all $\boldsymbol{\varphi} \in \mathcal{L}^2(t_1, T; \boldsymbol{V})$*, where*

$$\begin{aligned}
\langle \boldsymbol{r}_h, \boldsymbol{\varphi} \rangle = \langle \boldsymbol{f}_h^n - \boldsymbol{f}, \boldsymbol{\varphi} \rangle &+ \alpha\langle (\overline{\partial}^B - \overline{\partial}_n^B)\boldsymbol{m}_h^n, \boldsymbol{\varphi} \rangle + \langle (\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n) \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle \\
&+ (t - t_n)\langle (\overline{\partial}\boldsymbol{m}_h^n - \overline{\partial}\widehat{\boldsymbol{m}}_h^n) \times \overline{\partial}^B \boldsymbol{m}_h^n, \boldsymbol{\varphi} \rangle + (t - t_n)^2 \langle \overline{\partial}\boldsymbol{m}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n, \boldsymbol{\varphi} \rangle \\
&+ \langle (\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^n, \boldsymbol{\varphi} \rangle + (t - t_n)\langle (\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n)\overline{\partial}\lambda_h^n, \boldsymbol{\varphi} \rangle \\
&+ (t - t_n)^2 \langle \overline{\partial}\boldsymbol{m}_h^n \overline{\partial}\lambda_h^n, \boldsymbol{\varphi} \rangle + (t - t_n)\langle \boldsymbol{\Psi}, \boldsymbol{\varphi} \rangle \\
&+ \langle (\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n), \boldsymbol{\varphi} \rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n), \boldsymbol{\varphi} \rangle
\end{aligned} \tag{2.4.10}$$

for all $\boldsymbol{\varphi} \in \mathcal{L}^2(t_1, T; \boldsymbol{V})$ *and with*

$$\boldsymbol{\Psi} \coloneqq \alpha\overline{\partial}^2 \boldsymbol{m}_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n - \overline{\partial}(\Delta_h^n \boldsymbol{m}_h^n) + \widehat{\boldsymbol{m}}_h^n \overline{\partial}\lambda_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n \lambda_h^n. \tag{2.4.11}$$

Proof. We start by using the orthogonality of the $\mathcal{L}^2$ projection $\boldsymbol{P}_0^n$ onto $\boldsymbol{V}_h^n$, i.e., $\langle \boldsymbol{P}_0^n \boldsymbol{u}, \boldsymbol{\varphi} \rangle = \langle \boldsymbol{P}_0^n \boldsymbol{u}, \boldsymbol{P}_0^n \boldsymbol{\varphi} \rangle = \langle \boldsymbol{u}, \boldsymbol{P}_0^n \boldsymbol{\varphi} \rangle$ for $\boldsymbol{u}, \boldsymbol{\varphi} \in \boldsymbol{\mathcal{L}}^2(\Omega)$, the definition of the discrete Laplacian (2.2.2)

and the discrete equation (2.3.4), to obtain

$$\begin{aligned}
&\langle \alpha \overline{\partial}_n^B \boldsymbol{m}_h^n + \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n) - \Delta_h^n \boldsymbol{m}_h^n + \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \lambda_h^n) - \boldsymbol{f}_h^n, \boldsymbol{\varphi}\rangle \\
&\quad = \langle \alpha \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n - \Delta_h^n \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \lambda_h^n - \boldsymbol{f}_h^n, \boldsymbol{P}_0^n \boldsymbol{\varphi}\rangle \\
&\quad = \langle \alpha \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n - \boldsymbol{f}_h^n, \boldsymbol{P}_0^n \boldsymbol{\varphi}\rangle + a(\boldsymbol{m}_h^n, \boldsymbol{P}_0^n \boldsymbol{\varphi}) + b^{\widehat{\boldsymbol{m}}_h^n}(\lambda_h^n, \boldsymbol{P}_0^n \boldsymbol{\varphi}) = 0
\end{aligned} \tag{2.4.12}$$

for $n \geq 2$ and for all $\boldsymbol{\varphi} \in \mathcal{L}^2(t_1, T; \boldsymbol{V})$. By applying the weak formulation (2.3.3) and the definition of the error variables (2.4.8), we have for $\boldsymbol{\varphi} \in \mathcal{L}^2(t_1, T; \boldsymbol{V})$

$$\begin{aligned}
&\alpha\langle \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m} \times \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{\epsilon} \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi}) \\
&\quad = -\langle \boldsymbol{f}, \boldsymbol{\varphi}\rangle + \alpha\langle \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\boldsymbol{w}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}_h}(\lambda_h, \boldsymbol{\varphi}).
\end{aligned}$$

Subtracting equation (2.4.12) yields

$$\begin{aligned}
&\alpha\langle \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m} \times \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{\epsilon} \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi}) \\
&\quad = \langle \boldsymbol{f}_h^n - \boldsymbol{f}, \boldsymbol{\varphi}\rangle + \alpha\langle \partial_t \boldsymbol{M}_h - \overline{\partial}_n^B \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h \times \partial_t \boldsymbol{M}_h - \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n), \boldsymbol{\varphi}\rangle \\
&\qquad + a(\boldsymbol{w}, \boldsymbol{\varphi}) + \langle \Delta_h^n \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h \lambda_h - \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \lambda_h^n), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.4.13}$$

For the second term on the right-hand side, we use the definition of the three-point reconstruction (2.1.5) to get

$$\alpha\langle \partial_t \boldsymbol{M}_h - \overline{\partial}_n^B \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle = \alpha\langle (\overline{\partial}^B - \overline{\partial}_n^B)\boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle + (t - t_n)\alpha\langle \overline{\partial}^2 \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle. \tag{2.4.14}$$

Next, adding and subtracting $\langle \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle$ yields for the third term on the right-hand side

$$\begin{aligned}
\langle \boldsymbol{m}_h \times \partial_t \boldsymbol{M}_h - \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n), \boldsymbol{\varphi}\rangle &= \langle \boldsymbol{m}_h \times \partial_t \boldsymbol{M}_h - \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle \\
&\quad + \langle (\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.4.15}$$

Again, by adding and subtracting $\langle \widehat{\boldsymbol{m}}_h^n \lambda_h^n, \boldsymbol{\varphi}\rangle$ to the last term on the right-hand side, we obtain

$$\langle \boldsymbol{m}_h \lambda_h - \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \lambda_h^n), \boldsymbol{\varphi}\rangle = \langle \boldsymbol{m}_h \lambda_h - \widehat{\boldsymbol{m}}_h^n \lambda_h^n, \boldsymbol{\varphi}\rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n), \boldsymbol{\varphi}\rangle. \tag{2.4.16}$$

Further, we have, by the definition of $\boldsymbol{w}$ (2.4.8), as well as the definition of the elliptic reconstruction (2.2.3)

$$\begin{aligned}
a(\boldsymbol{w}, \boldsymbol{\varphi}) &= a(\boldsymbol{\mathcal{R}} \boldsymbol{m}_h^n, \boldsymbol{\varphi}) + (t - t_n)\, a(\overline{\partial} \boldsymbol{\mathcal{R}} \boldsymbol{m}_h^n, \boldsymbol{\varphi}) \\
&= -\langle \Delta_h^n \boldsymbol{m}_h^n, \boldsymbol{\varphi}\rangle - (t - t_n)\langle \overline{\partial}(\Delta_h^n \boldsymbol{m}_h^n), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.4.17}$$

Utilizing (2.4.14)–(2.4.17) to rewrite the right-hand side of (2.4.13), we obtain

$$\begin{aligned}
&\alpha\langle\partial_t\boldsymbol{\mathcal{E}},\boldsymbol{\varphi}\rangle + \langle\boldsymbol{m}\times\partial_t\boldsymbol{\mathcal{E}},\boldsymbol{\varphi}\rangle + \langle\boldsymbol{\epsilon}\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}},\boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon,\boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h,\boldsymbol{\varphi})\\
&\quad= \langle\boldsymbol{f}_h^n-\boldsymbol{f},\boldsymbol{\varphi}\rangle + \alpha\langle(\overline{\partial}^B-\overline{\partial}_n^B)\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle + \langle\boldsymbol{m}_h\times\partial_t\boldsymbol{M}_h - \widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle\\
&\qquad+\langle\boldsymbol{m}_h\lambda_h-\widehat{\boldsymbol{m}}_h^n\lambda_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)\Big(\alpha\langle\overline{\partial}^2\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle - \langle\overline{\partial}(\Delta_h^n\boldsymbol{m}_h^n),\boldsymbol{\varphi}\rangle\Big)\\
&\qquad+\langle(\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n),\boldsymbol{\varphi}\rangle + \langle(\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\lambda_h^n),\boldsymbol{\varphi}\rangle.
\end{aligned}\tag{2.4.18}$$

Next, we continue with rewriting the right-hand side of (2.4.18). We first add and subtract $(t-t_n)\langle\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle$, and second $(t-t_n)\langle\overline{\partial}\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle$. With the definition of the three-point reconstruction $\boldsymbol{M}_h$ (2.1.5) and the linear reconstruction $\boldsymbol{m}_h$ (2.1.4), this yields for the third term on the right-hand side of (2.4.18)

$$\begin{aligned}
&\langle\boldsymbol{m}_h\times\partial_t\boldsymbol{M}_h - \widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle\\
&\quad=\langle(\boldsymbol{m}_h^n-\widehat{\boldsymbol{m}}_h^n)\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle + (t-t_n)\langle\overline{\partial}\boldsymbol{m}_h^n\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle + (t-t_n)\langle\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle\\
&\quad=\langle(\boldsymbol{m}_h^n-\widehat{\boldsymbol{m}}_h^n)\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle + (t-t_n)\langle(\overline{\partial}\boldsymbol{m}_h^n-\overline{\partial}\widehat{\boldsymbol{m}}_h^n)\times\overline{\partial}^B\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle\\
&\qquad+(t-t_n)^2\langle\overline{\partial}\boldsymbol{m}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)\langle\overline{\partial}\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n+\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle.
\end{aligned}\tag{2.4.19}$$

Similarly, by first adding and subtracting $\langle\widehat{\boldsymbol{m}}_h\lambda_h^n,\boldsymbol{\varphi}\rangle$, and second $(t-t_n)\langle\widehat{\boldsymbol{m}}_h^n\overline{\partial}\lambda_h^n,\boldsymbol{\varphi}\rangle$ and using the definition of the linear reconstruction (2.1.4), we rewrite the fourth term on the right-hand side of (2.4.18) by

$$\begin{aligned}
\langle\boldsymbol{m}_h\lambda_h-\widehat{\boldsymbol{m}}_h^n\lambda_h^n,\boldsymbol{\varphi}\rangle &= \langle(\boldsymbol{m}_h-\widehat{\boldsymbol{m}}_h)\lambda_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)\langle\boldsymbol{m}_h\overline{\partial}\lambda_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)\langle\overline{\partial}\widehat{\boldsymbol{m}}_h^n\lambda_h^n,\boldsymbol{\varphi}\rangle\\
&=\langle(\boldsymbol{m}_h-\widehat{\boldsymbol{m}}_h)\lambda_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)\langle(\boldsymbol{m}_h^n-\widehat{\boldsymbol{m}}_h^n)\overline{\partial}\lambda_h^n,\boldsymbol{\varphi}\rangle\\
&\quad+(t-t_n)\langle\widehat{\boldsymbol{m}}_h^n\overline{\partial}\lambda_h^n+\overline{\partial}\widehat{\boldsymbol{m}}_h^n\lambda_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)^2\langle\overline{\partial}\boldsymbol{m}_h^n\overline{\partial}\lambda_h^n,\boldsymbol{\varphi}\rangle.
\end{aligned}\tag{2.4.20}$$

Finally, by inserting (2.4.19) and (2.4.20) into the right-hand side of (2.4.18), we conclude

$$\begin{aligned}
&\alpha\langle\partial_t\boldsymbol{\mathcal{E}},\boldsymbol{\varphi}\rangle + \langle\boldsymbol{m}\times\partial_t\boldsymbol{\mathcal{E}},\boldsymbol{\varphi}\rangle + \langle\boldsymbol{\epsilon}\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}},\boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon,\boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h,\boldsymbol{\varphi})\\
&\quad=\langle\boldsymbol{f}_h^n-\boldsymbol{f},\boldsymbol{\varphi}\rangle + \alpha\langle(\overline{\partial}^B-\overline{\partial}_n^B)\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle + \langle(\boldsymbol{m}_h^n-\widehat{\boldsymbol{m}}_h^n)\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle\\
&\qquad+(t-t_n)\langle(\overline{\partial}\boldsymbol{m}_h^n-\overline{\partial}\widehat{\boldsymbol{m}}_h^n)\times\overline{\partial}^B\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)^2\langle\overline{\partial}\boldsymbol{m}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle\\
&\qquad+\langle(\boldsymbol{m}_h-\widehat{\boldsymbol{m}}_h)\lambda_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)\langle(\boldsymbol{m}_h^n-\widehat{\boldsymbol{m}}_h^n)\overline{\partial}\lambda_h^n,\boldsymbol{\varphi}\rangle\\
&\qquad+(t-t_n)^2\langle\overline{\partial}\boldsymbol{m}_h^n\overline{\partial}\lambda_h^n,\boldsymbol{\varphi}\rangle + (t-t_n)\langle\boldsymbol{\Psi},\boldsymbol{\varphi}\rangle\\
&\qquad+\langle(\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n),\boldsymbol{\varphi}\rangle + \langle(\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\lambda_h^n),\boldsymbol{\varphi}\rangle,
\end{aligned}$$

where $\boldsymbol{\Psi}$ is defined in (2.4.11). □

Having established the parabolic error equation, we can now proceed with the proof of Theorem 19. For clarity and to emphasize the key ideas, we postpone the proofs of the a posteriori terms and the initial time step a posteriori bound to the next section.

Proof of Theorem 19. Rewriting the parabolic error equation (2.4.9), we have

$$\begin{aligned}\alpha\langle\partial_t\widetilde{\boldsymbol{\mathcal{E}}},\boldsymbol{\varphi}\rangle + \langle\boldsymbol{m}\times\partial_t\widetilde{\boldsymbol{\mathcal{E}}},\boldsymbol{\varphi}\rangle + \langle\widetilde{\boldsymbol{\mathcal{E}}}\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}},\boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon,\boldsymbol{\varphi}) + b^{\widetilde{\boldsymbol{\mathcal{E}}}}(\lambda_h,\boldsymbol{\varphi})\\ = \alpha\langle\partial_t(\boldsymbol{W}-\boldsymbol{M}_h),\boldsymbol{\varphi}\rangle + \langle\boldsymbol{m}\times\partial_t(\boldsymbol{W}-\boldsymbol{M}_h),\boldsymbol{\varphi}\rangle + \langle(\boldsymbol{W}-\boldsymbol{m}_h)\times\partial_t\boldsymbol{M}_h,\boldsymbol{\varphi}\rangle\\ + \langle(\boldsymbol{W}-\boldsymbol{m}_h)\lambda_h,\boldsymbol{\varphi}\rangle + \langle\boldsymbol{r}_h,\boldsymbol{\varphi}\rangle.\end{aligned}$$

Now, testing with $\boldsymbol{\varphi}=\partial_t\widetilde{\boldsymbol{\mathcal{E}}}$ yields

$$\begin{aligned}\alpha\|\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + a(\widetilde{\boldsymbol{\epsilon}},\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) + b^{\boldsymbol{m}}(\varepsilon,\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) + b^{\widetilde{\boldsymbol{\mathcal{E}}}}(\lambda_h,\partial_t\widetilde{\boldsymbol{\mathcal{E}}})\\ \le \|\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\|\Big(\|\partial_t\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{L}}^\infty}\|\widetilde{\boldsymbol{\mathcal{E}}}\| + (1+\alpha)\|\partial_t(\boldsymbol{W}-\boldsymbol{M}_h)\|\\ + \|(\boldsymbol{W}-\boldsymbol{m}_h)\times\partial_t\boldsymbol{M}_h\| + \|(\boldsymbol{W}-\boldsymbol{m}_h)\lambda_h\| + \|\boldsymbol{r}_h\|\Big)\end{aligned}$$

since $\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{L}}^\infty}=1$ and $\langle\boldsymbol{m}\times\partial_t\widetilde{\boldsymbol{\mathcal{E}}},\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\rangle=0$. Next, we treat the bilinear form a as

$$a(\widetilde{\boldsymbol{\epsilon}},\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) = a(\widetilde{\boldsymbol{\mathcal{E}}},\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) + a(\boldsymbol{w}-\boldsymbol{W},\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) = a(\widetilde{\boldsymbol{\mathcal{E}}},\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) + \langle-\Delta_h(\boldsymbol{w}-\boldsymbol{W}),\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\rangle,$$

where $-\Delta_h(\boldsymbol{w}-\boldsymbol{W}) = \frac{1}{2}(t-t_n)(t-t_{n-1})\overline{\partial}^2(-\Delta_h^n)\boldsymbol{m}_h^n$ by the definition (2.4.5) and (2.4.6) of the reconstructions for $t\in I_n$, since $\boldsymbol{w}$ and $\boldsymbol{W}$ are defined through pointwise (in time) elliptic reconstructions. Further, using $a(\widetilde{\boldsymbol{\mathcal{E}}},\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\nabla}\widetilde{\boldsymbol{\mathcal{E}}}\|^2$, we obtain

$$\begin{aligned}\alpha\|\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\nabla}\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + b^{\boldsymbol{m}}(\varepsilon,\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) + b^{\widetilde{\boldsymbol{\mathcal{E}}}}(\lambda_h,\partial_t\widetilde{\boldsymbol{\mathcal{E}}})\\ \le \|\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\|\Big(\|\partial_t\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{L}}^\infty}\|\widetilde{\boldsymbol{\mathcal{E}}}\| + (1+\alpha)\|\partial_t(\boldsymbol{W}-\boldsymbol{M}_h)\| + \|(\boldsymbol{W}-\boldsymbol{m}_h)\times\partial_t\boldsymbol{M}_h\|\\ + \|(\boldsymbol{W}-\boldsymbol{m}_h)\lambda_h\| + \|\boldsymbol{r}_h\| + \|\Delta_h(\boldsymbol{w}-\boldsymbol{W})\|\Big)\end{aligned}\tag{2.4.21}$$

To reduce the notation we define $c_\infty := \|\partial_t\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\lambda_h\|_{\mathcal{L}^\infty}$, as well as,

$$\mathcal{A} := (1+\alpha)\|\partial_t(\boldsymbol{W}-\boldsymbol{M}_h)\| + c_\infty\|(\boldsymbol{W}-\boldsymbol{m}_h)\| + \|\boldsymbol{r}_h\| + \|\Delta_h(\boldsymbol{w}-\boldsymbol{W})\|,\tag{2.4.22}$$

Thus we can rewrite equation (2.4.21) by

$$\alpha\|\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\nabla}\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + b^{\boldsymbol{m}}(\varepsilon,\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) + b^{\widetilde{\boldsymbol{\mathcal{E}}}}(\lambda_h,\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) \le \|\partial_t\widetilde{\boldsymbol{\mathcal{E}}}\|\Big(\|\partial_t\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{L}}^\infty}\|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A}\Big).$$

Further, using the parabolic error equation (2.4.9), $\boldsymbol{m}\cdot\partial_t\boldsymbol{m}=0$ a.e. and $\boldsymbol{m}\cdot\boldsymbol{P}(\boldsymbol{m})\partial_t\boldsymbol{M}_h=0$ a.e., we obtain

$$\begin{aligned}b^{\boldsymbol{m}}(\varepsilon,\partial_t\widetilde{\boldsymbol{\mathcal{E}}}) = b^{\boldsymbol{m}}(\varepsilon,\partial_t\boldsymbol{W}) = b^{\boldsymbol{m}}(\varepsilon,(\boldsymbol{I}-\boldsymbol{P}(\boldsymbol{m}))\partial_t\boldsymbol{W})\\ \le \|(\boldsymbol{I}-\boldsymbol{P}(\boldsymbol{m}))\partial_t\boldsymbol{W}\|\Big((1+\alpha)\|\partial_t\boldsymbol{\mathcal{E}}\| + \|\boldsymbol{\epsilon}\|\|\partial_t\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{L}}^\infty} + \|\boldsymbol{\epsilon}\|\|\lambda_h\|_{\mathcal{L}^\infty} + \|\boldsymbol{r}_h\|\Big)\\ - a(\widetilde{\boldsymbol{\epsilon}},(\boldsymbol{I}-\boldsymbol{P}(\boldsymbol{m}))\partial_t\boldsymbol{W})\\ \le \|(\boldsymbol{I}-\boldsymbol{P}(\boldsymbol{m}))\partial_t\boldsymbol{W}\|\Big((1+\alpha)\|\partial_t\boldsymbol{\mathcal{E}}\| + c_\infty\|\boldsymbol{\epsilon}\| + \|\boldsymbol{r}_h\| + \|\Delta_h(\boldsymbol{w}-\boldsymbol{W})\|\Big)\\ + \|\boldsymbol{\nabla}\widetilde{\boldsymbol{\mathcal{E}}}\|\|\boldsymbol{\nabla}(\boldsymbol{I}-\boldsymbol{P}(\boldsymbol{m}))\partial_t\boldsymbol{W}\|,\end{aligned}$$

since

$$a(\widetilde{\boldsymbol{\epsilon}}, (\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}) = a(\widetilde{\boldsymbol{\mathcal{E}}}, (\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}) + \langle -\Delta_h(\boldsymbol{w} - \boldsymbol{W}), (\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\rangle.$$

Using $\|\boldsymbol{\epsilon}\| \leq \|\widetilde{\boldsymbol{\mathcal{E}}}\| + \|\boldsymbol{W} - \boldsymbol{m}_h\|$, $\|\partial_t \boldsymbol{\mathcal{E}}\| \leq \|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\| + \|\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|$ and the definition of $\mathcal{A}$ (2.4.22) yields

$$\begin{aligned} b^{\boldsymbol{m}}(\varepsilon, \partial_t \widetilde{\boldsymbol{\mathcal{E}}}) \leq \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| \Big((1+\alpha)\|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\| + c_\infty \|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A} \Big) \\ + \|\boldsymbol{\nabla} \widetilde{\boldsymbol{\mathcal{E}}}\| \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\|. \end{aligned} \tag{2.4.23}$$

Utilizing $b^{\widetilde{\boldsymbol{\mathcal{E}}}}(\lambda_h, \partial_t \widetilde{\boldsymbol{\mathcal{E}}}) \leq \|\lambda_h\|_{\mathcal{L}^\infty} \|\widetilde{\boldsymbol{\mathcal{E}}}\| \|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\|$ and (2.4.23), we deduce

$$\begin{aligned} \alpha \|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\nabla} \widetilde{\boldsymbol{\mathcal{E}}}\|^2 \leq \|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\| \Big(c_\infty \|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A} \Big) + \|\boldsymbol{\nabla} \widetilde{\boldsymbol{\mathcal{E}}}\| \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| \\ + \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| \Big((1+\alpha)\|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\| + c_\infty \|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A} \Big). \end{aligned}$$

Applying Young's inequality, making use of

$$\|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\|^2 \geq \frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\boldsymbol{\mathcal{E}}}\|^2 - \|\widetilde{\boldsymbol{\mathcal{E}}}\|^2$$

and subtracting $\frac{1}{2}\alpha \|\partial_t \widetilde{\boldsymbol{\mathcal{E}}}\|^2$ from both sides, we obtain

$$\begin{aligned} \min\{\alpha, 1\}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\boldsymbol{\mathcal{E}}}\|_{\boldsymbol{\mathcal{H}}^1}^2 \leq \|\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + \frac{2}{\alpha}\Big(c_\infty \|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A} \Big)^2 + 2\|\boldsymbol{\nabla} \widetilde{\boldsymbol{\mathcal{E}}}\| \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| \\ + \frac{2(1+\alpha)^2}{\alpha}\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\|^2 + 2\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| \Big(c_\infty \|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A} \Big). \end{aligned}$$

Define

$$\gamma_{0,\boldsymbol{W}} \coloneqq (1 + \|\boldsymbol{W}\|_{\boldsymbol{\mathcal{L}}^\infty})\|\partial_t \boldsymbol{W}\|_{\boldsymbol{\mathcal{L}}^\infty}, \qquad \gamma_{1,\boldsymbol{W}} \coloneqq \gamma(\boldsymbol{W}, \partial_t \boldsymbol{W}), \tag{2.4.24}$$

where γ is defined in Lemma 16. Next, we estimate using the Lipschitz-type bound of the orthogonal projection in Lemma 16

$$\begin{aligned} \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| &\leq \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\| + \|(\boldsymbol{P}(\boldsymbol{W}) - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| \\ &\leq \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\| + \gamma_{0,\boldsymbol{W}}\|\widetilde{\boldsymbol{\mathcal{E}}}\|, \\ \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| &\leq \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\| + \|\boldsymbol{\nabla}(\boldsymbol{P}(\boldsymbol{W}) - \boldsymbol{P}(\boldsymbol{m}))\partial_t \boldsymbol{W}\| \\ &\leq \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\| + \gamma_{1,\boldsymbol{W}}\|\widetilde{\boldsymbol{\mathcal{E}}}\|_{\boldsymbol{\mathcal{H}}^1}. \end{aligned}$$

Applying the above equations yields

$$\begin{aligned} \min\{\alpha, 1\}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\boldsymbol{\mathcal{E}}}\|_{\boldsymbol{\mathcal{H}}^1}^2 \leq \|\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + \frac{2}{\alpha}\Big(c_\infty \|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A} \Big)^2 + 2\|\boldsymbol{\nabla} \widetilde{\boldsymbol{\mathcal{E}}}\| \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\| \\ + \frac{2(1+\alpha)^2}{\alpha}\Big(\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\| + \gamma_{0,\boldsymbol{W}}\|\widetilde{\boldsymbol{\mathcal{E}}}\| \Big)^2 \end{aligned}$$

$$+ 2(\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\| + \gamma_{0,\boldsymbol{W}}\|\widetilde{\boldsymbol{\mathcal{E}}}\|)\Big(c_\infty\|\widetilde{\boldsymbol{\mathcal{E}}}\| + \mathcal{A}\Big)$$
$$+ 2\|\boldsymbol{\nabla}\widetilde{\boldsymbol{\mathcal{E}}}\|\gamma_{1,\boldsymbol{W}}\|\widetilde{\boldsymbol{\mathcal{E}}}\|_{\boldsymbol{\mathcal{H}}^1}.$$

Exploiting Young's inequality, $(a+b)^2 \le 2(a^2+b^2)$ and reordering yields

$$\min\{\alpha,1\}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{\boldsymbol{\mathcal{E}}}\|^2_{\boldsymbol{\mathcal{H}}^1} \le c_0\|\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + c_1\|\widetilde{\boldsymbol{\mathcal{E}}}\|^2_{\boldsymbol{\mathcal{H}}^1} + \Big(\frac{4(1+\alpha)^2}{\alpha} + 1 + c_\infty^2\Big)\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t\boldsymbol{W}\|^2$$
$$+ \Big(\frac{4}{\alpha} + 2\Big)\mathcal{A}^2 + \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t\boldsymbol{W}\|^2,$$

where $c_0 \coloneqq 2 + \frac{4}{\alpha}c_\infty^2 + 2c_\infty\gamma_{0,\boldsymbol{W}} + (\frac{4(1+\alpha)^2}{\alpha} + 1)\gamma_{0,\boldsymbol{W}}^2$ and $c_1 \coloneqq 1 + 2\gamma_{1,\boldsymbol{W}}$. Integrating from $t = t_1$ to t_N, we obtain

$$\|\widetilde{\boldsymbol{\mathcal{E}}}(t_N)\|^2_{\boldsymbol{\mathcal{H}}^1} \le \|\widetilde{\boldsymbol{\mathcal{E}}}(t_1)\|^2_{\boldsymbol{\mathcal{H}}^1} + \frac{1}{\min\{\alpha,1\}}\int_{t_1}^{t_N}\Big(c_0\|\widetilde{\boldsymbol{\mathcal{E}}}\|^2 + c_1\|\widetilde{\boldsymbol{\mathcal{E}}}\|^2_{\boldsymbol{\mathcal{H}}^1}\Big)\,\mathrm{d}t + \frac{1}{\min\{\alpha,1\}}\widetilde{\mathcal{A}},$$

where

$$\widetilde{\mathcal{A}} \coloneqq \int_{t_1}^{t_N}\Big\{\widetilde{c}_\infty\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t\boldsymbol{W}\|^2 + \Big(\frac{4}{\alpha} + 2\Big)\mathcal{A}^2 + \|\boldsymbol{\nabla}(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t\boldsymbol{W}\|^2\Big\}\,\mathrm{d}t$$

with $\widetilde{c}_\infty \coloneqq \frac{4(1+\alpha)^2}{\alpha} + 1 + c_\infty^2$. Thus, estimating $\|\widetilde{\boldsymbol{\mathcal{E}}}\|^2 \le \|\widetilde{\boldsymbol{\mathcal{E}}}\|^2_{\boldsymbol{\mathcal{H}}^1}$, we conclude by Gronwall's inequality

$$\|\widetilde{\boldsymbol{\mathcal{E}}}(t_N)\|^2_{\boldsymbol{\mathcal{H}}^1} \le \Big(\|\widetilde{\boldsymbol{\mathcal{E}}}(t_1)\|^2_{\boldsymbol{\mathcal{H}}^1} + \frac{1}{\min\{\alpha,1\}}\widetilde{\mathcal{A}}\Big)\mathrm{e}^{\int_{t_1}^{t_N}\frac{(c_0+c_1)}{\min\{\alpha,1\}}\,\mathrm{d}t}.$$

We finalize the a posteriori estimate (2.4.4) for $t \in [t_1, t_N]$ by

$$\|\boldsymbol{\nabla}(\boldsymbol{m} - \boldsymbol{m}_h)(t)\|^2 \le \|\boldsymbol{\nabla}(\boldsymbol{M}_h - \boldsymbol{m}_h)(t)\|^2 + \|\boldsymbol{\nabla}(\boldsymbol{W} - \boldsymbol{M}_h)(t)\|^2 + \|\boldsymbol{\nabla}\widetilde{\boldsymbol{\mathcal{E}}}(t)\|^2,$$

where we estimate the remaining terms in the next section. □

2.5 Remaining error estimates

In this section, we provide the remaining a posteriori error bounds to conclude Theorem 19. The error estimates are mostly derived by adapting the approach of [25], where comparable error terms were analyzed. We start by taking care of the initial time-step, which is provided by (2.3.7). To simplify the notation, we assume $\boldsymbol{V}_h^0 = \boldsymbol{V}_h^1 = \boldsymbol{V}_h^2$.

Lemma 21 (Initial time-step a posteriori estimate)**.** *Assume $V_h^0 = V_h^1 = V_h^2$. Let $(\boldsymbol{m}, \lambda)$ be the exact solution of (2.3.3). Further, let $(\boldsymbol{m}_h^1, \lambda_h^1)$ be the solution of (2.3.7) and $(\boldsymbol{m}_h^2, \lambda_h^2)$ be the solution of (2.3.4). Then*

$$\begin{aligned}
&\|\boldsymbol{\nabla}(\boldsymbol{m} - \boldsymbol{m}_h)\|^2_{\mathcal{L}^\infty(0,t_1;\boldsymbol{\mathcal{L}}^2(\Omega))} \\
&\quad\lesssim \|\boldsymbol{\nabla}(\boldsymbol{P}_0^0\boldsymbol{m}(0) - \boldsymbol{m}(0))\|^2 + \|\boldsymbol{\nabla}(\boldsymbol{M}_h - \boldsymbol{m}_h)\|^2_{\mathcal{L}^\infty(0,t_1;\boldsymbol{\mathcal{L}}^2(\Omega))} + \|\boldsymbol{\nabla}(\boldsymbol{W} - \boldsymbol{M}_h)\|^2_{\mathcal{L}^\infty(0,t_1;\boldsymbol{\mathcal{L}}^2(\Omega))} \\
&\qquad + \int_0^{t_1} \Big(\|\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|^2 + \|\boldsymbol{W} - \boldsymbol{m}_h\|^2 + \|\boldsymbol{r}_h^0\|^2 + \|\Delta_h(\boldsymbol{w} - \boldsymbol{W})\|^2 \\
&\qquad\qquad + \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\|^2_{\boldsymbol{\mathcal{H}}^1} \Big)\, \mathrm{d}t,
\end{aligned}$$

where the initial reconstruction $\boldsymbol{M}_h$ is given in (2.1.7) and $\boldsymbol{r}_h^0$ is defined by

$$\begin{aligned}
\langle \boldsymbol{r}_h^0, \boldsymbol{\varphi}\rangle &= \langle \boldsymbol{f}_h - \boldsymbol{f}, \boldsymbol{\varphi}\rangle + \langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + (t - t_1)\langle (\overline{\partial}\boldsymbol{m}_h^1 - \overline{\partial}\widehat{\boldsymbol{m}}_h^1) \times \overline{\partial}\boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle \\
&\quad + (t - t_1)(t - t_{1/2})\langle \overline{\partial}\boldsymbol{m}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^2, \boldsymbol{\varphi}\rangle + \langle (\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^1, \boldsymbol{\varphi}\rangle \\
&\quad + (t - t_1)\langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1)\overline{\partial}\lambda_h^1, \boldsymbol{\varphi}\rangle - \frac{\tau^2}{4}\langle \overline{\partial}\widehat{\boldsymbol{m}}_h^1 \overline{\partial}\lambda_h^1, \boldsymbol{\varphi}\rangle + (t - t_1)^2\langle \overline{\partial}\boldsymbol{m}_h^1 \overline{\partial}\lambda_h^1, \boldsymbol{\varphi}\rangle \\
&\quad + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial}\boldsymbol{m}_h^1), \boldsymbol{\varphi}\rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}), \boldsymbol{\varphi}\rangle + (t - t_{1/2})\langle \boldsymbol{\Psi}_0, \boldsymbol{\varphi}\rangle
\end{aligned} \tag{2.5.1}$$

with $\boldsymbol{\Psi}_0 = \alpha\overline{\partial}^2\boldsymbol{m}_h^2 + \overline{\partial}\widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}\boldsymbol{m}_h^1 + \widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}^2\boldsymbol{m}_h^2 - \overline{\partial}(\Delta_h^1\boldsymbol{m}_h^1) + \widehat{\boldsymbol{m}}_h^1\overline{\partial}\lambda_h^1 + \overline{\partial}\widehat{\boldsymbol{m}}_h^1\lambda_h^1 - \overline{\partial}\boldsymbol{f}_h^1$.

Proof. Since $\boldsymbol{V}_h^1 = \boldsymbol{V}_h^0$, we have $\overline{\partial}\boldsymbol{m}_h^1 = \boldsymbol{P}_0^1\overline{\partial}\boldsymbol{m}_h^1 = \overline{\partial}_1\boldsymbol{m}_h^1$. Similarly to (2.4.12), using the orthogonality of the $\mathcal{L}^2$ projection $\boldsymbol{P}_0^1$ onto $\boldsymbol{V}_h^1 = \boldsymbol{V}_h^0$, the definition of the discrete Laplacian (2.2.2) and the weak formulation of the trapezoidal scheme (2.3.7) we obtain

$$\begin{aligned}
&\langle \alpha\overline{\partial}\boldsymbol{m}_h^1 + \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial}\boldsymbol{m}_h^1) - \Delta_h^1\boldsymbol{m}_h^{1/2} + \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}) - \boldsymbol{f}_h^{1/2}, \boldsymbol{\varphi}\rangle \\
&\quad = \langle \alpha\overline{\partial}\boldsymbol{m}_h^1 + \widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial}\boldsymbol{m}_h^1 - \Delta_h^1\boldsymbol{m}_h^{1/2} + \widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2} - \boldsymbol{f}_h^{1/2}, \boldsymbol{P}_0^1\boldsymbol{\varphi}\rangle \\
&\quad = \langle \alpha\overline{\partial}\boldsymbol{m}_h^1 + \widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial}\boldsymbol{m}_h^1 - \boldsymbol{f}_h^{1/2}, \boldsymbol{P}_0^1\boldsymbol{\varphi}\rangle + a(\boldsymbol{m}_h^{1/2}, \boldsymbol{P}_0^1\boldsymbol{\varphi}) + b^{\widehat{\boldsymbol{m}}_h^{1/2}}(\lambda_h^{1/2}, \boldsymbol{P}_0^1\boldsymbol{\varphi}) = 0
\end{aligned} \tag{2.5.2}$$

for all $\boldsymbol{\varphi} \in \mathcal{L}^2(0, t_1; \boldsymbol{V})$. Thus, by applying the weak formulation (2.3.3) and the definition of the error variables (2.4.8), we have for $\boldsymbol{\varphi} \in \mathcal{L}^2(0, t_1; \boldsymbol{V})$

$$\begin{aligned}
&\alpha\langle \partial_t\boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m} \times \partial_t\boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{\epsilon} \times \partial_t\boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi}) \\
&\quad = -\langle \boldsymbol{f}, \boldsymbol{\varphi}\rangle + \alpha\langle \partial_t\boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h \times \partial_t\boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\boldsymbol{w}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}_h}(\lambda_h, \boldsymbol{\varphi}).
\end{aligned}$$

Subtracting equation (2.5.2) and adding $\langle \widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial}\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial}\boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle = 0$, as well as

$\langle \widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2} - \widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}, \boldsymbol{\varphi}\rangle = 0$, yields

$$\begin{aligned}
&\alpha\langle \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}\times\partial_t\boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{\epsilon}\times\partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi})\\
&\quad = \langle \boldsymbol{f}_h^{1/2} - \boldsymbol{f}, \boldsymbol{\varphi}\rangle + \alpha\langle \partial_t \boldsymbol{M}_h - \overline{\partial}\boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h\times\partial_t\boldsymbol{M}_h - \widehat{\boldsymbol{m}}_h^{1/2}\times\overline{\partial}\boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle\\
&\qquad a(\boldsymbol{w}, \boldsymbol{\varphi}) + \langle \Delta_h^1 \boldsymbol{m}_h^{1/2}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h\lambda_h - \widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}, \boldsymbol{\varphi}\rangle\\
&\qquad + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\times\overline{\partial}\boldsymbol{m}_h^1), \boldsymbol{\varphi}\rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.5.3}$$

For the second term on the right-hand side, we use the definition of the three-point reconstruction $\boldsymbol{M}_h$ on I_1 (2.1.7) to obtain

$$\alpha\langle \partial_t \boldsymbol{M}_h - \overline{\partial}\boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle = \alpha(t - t_{1/2})\langle \overline{\partial}^2 \boldsymbol{m}_h^2, \boldsymbol{\varphi}\rangle. \tag{2.5.4}$$

Further, we have, by the definition of $\boldsymbol{w}$ (2.4.5), as well as the definition of the elliptic reconstruction (2.2.3)

$$\begin{aligned}
a(\boldsymbol{w}, \boldsymbol{\varphi}) &= a(\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^1, \boldsymbol{\varphi}) + (t - t_1)\, a(\overline{\partial}\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^1, \boldsymbol{\varphi})\\
&= -\langle \Delta_h^1 \boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle - (t - t_1)\langle \overline{\partial}(\Delta_h^1 \boldsymbol{m}_h^1), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.5.5}$$

Substituting (2.5.4) and (2.5.5) into the right-hand side of (2.5.3) leads to

$$\begin{aligned}
&\alpha\langle \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}\times\partial_t\boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{\epsilon}\times\partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi})\\
&\quad = \langle \boldsymbol{f}_h^{1/2} - \boldsymbol{f}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h\times\partial_t\boldsymbol{M}_h - \widehat{\boldsymbol{m}}_h^{1/2}\times\overline{\partial}\boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle + \langle \Delta_h^1 \boldsymbol{m}_h^{1/2} - \Delta_h^1 \boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle\\
&\qquad + \langle \boldsymbol{m}_h\lambda_h - \widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}, \boldsymbol{\varphi}\rangle + \alpha(t - t_{1/2})\langle \overline{\partial}^2 \boldsymbol{m}_h^2, \boldsymbol{\varphi}\rangle - (t - t_1)\langle \overline{\partial}(\Delta_h^1 \boldsymbol{m}_h^1), \boldsymbol{\varphi}\rangle\\
&\qquad + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\times\overline{\partial}\boldsymbol{m}_h^1), \boldsymbol{\varphi}\rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.5.6}$$

We continue with rewriting the right-hand side of (2.5.6). Since $t - t_{1/2} = \frac{\tau}{2} + t - t_1$, we have

$$\boldsymbol{f}_h^1 - \boldsymbol{f}_h^{1/2} = \frac{\tau}{2}\overline{\partial}\boldsymbol{f}_h^1 = (t - t_{1/2})\overline{\partial}\boldsymbol{f}_h^1 - (t - t_1)\overline{\partial}\boldsymbol{f}_h^1 \tag{2.5.7}$$

The same identity holds for $\widehat{\boldsymbol{m}}_h^1$, λ_h^1 and $\Delta_h^1\boldsymbol{m}_h^1$. Inserting (2.5.7) for $\boldsymbol{f}_h^1$ and $\Delta_h^1\boldsymbol{m}_h^1$ into the right-hand side of (2.5.6) yields

$$\begin{aligned}
&\alpha\langle \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}\times\partial_t\boldsymbol{\mathcal{E}}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{\epsilon}\times\partial_t \boldsymbol{M}_h, \boldsymbol{\varphi}\rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi})\\
&\quad = \langle \boldsymbol{f}_h - \boldsymbol{f}, \boldsymbol{\varphi}\rangle + \langle \boldsymbol{m}_h\times\partial_t\boldsymbol{M}_h - \widehat{\boldsymbol{m}}_h^{1/2}\times\overline{\partial}\boldsymbol{m}_h^1, \boldsymbol{\varphi}\rangle\\
&\qquad + \langle \boldsymbol{m}_h\lambda_h - \widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}, \boldsymbol{\varphi}\rangle + (t - t_{1/2})\langle \alpha\overline{\partial}^2 \boldsymbol{m}_h^2 - \overline{\partial}(\Delta_h^1 \boldsymbol{m}_h^1) - \overline{\partial}\boldsymbol{f}_h^1, \boldsymbol{\varphi}\rangle\\
&\qquad + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\times\overline{\partial}\boldsymbol{m}_h^1), \boldsymbol{\varphi}\rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2}), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.5.8}$$

By adding and subtracting $\langle \widehat{\boldsymbol{m}}_h^1\times\partial_t\boldsymbol{M}_h, \boldsymbol{\varphi}\rangle$ to the second term on the right-hand side of

(2.5.8), the definition of the three-point reconstruction $\boldsymbol{M}_h$ on I_1 (2.1.7) and (2.5.7), we obtain

$$\begin{aligned}
&\langle \boldsymbol{m}_h \times \partial_t \boldsymbol{M}_h - \widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial} \boldsymbol{m}_h^1, \boldsymbol{\varphi} \rangle \\
&\quad = \langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle + (t - t_1) \langle \overline{\partial} \boldsymbol{m}_h^1 \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle + \langle (\widehat{\boldsymbol{m}}_h^1 - \widehat{\boldsymbol{m}}_h^{1/2}) \times \overline{\partial} \boldsymbol{m}_h^1, \boldsymbol{\varphi} \rangle \\
&\qquad + (t - t_{1/2}) \langle \widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^2, \boldsymbol{\varphi} \rangle \\
&\quad = \langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle + (t - t_1) \langle (\overline{\partial} \boldsymbol{m}_h^1 - \overline{\partial} \widehat{\boldsymbol{m}}_h^1) \times \overline{\partial} \boldsymbol{m}_h^1, \boldsymbol{\varphi} \rangle \\
&\qquad + (t - t_1)(t - t_{1/2}) \langle \overline{\partial} \boldsymbol{m}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^2, \boldsymbol{\varphi} \rangle + (t - t_{1/2}) \langle \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \times \overline{\partial} \boldsymbol{m}_h^1 + \widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^2, \boldsymbol{\varphi} \rangle.
\end{aligned} \tag{2.5.9}$$

Similarly, adding and subtracting first $\langle \widehat{\boldsymbol{m}}_h \lambda_h^1, \boldsymbol{\varphi} \rangle$ and second $(t - t_1) \langle \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle$ to the third term of the right-hand side (2.5.8) yields

$$\begin{aligned}
&\langle \boldsymbol{m}_h \lambda_h - \widehat{\boldsymbol{m}}_h^{1/2} \lambda_h^{1/2}, \boldsymbol{\varphi} \rangle \\
&\qquad = \langle (\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h) \lambda_h^1, \boldsymbol{\varphi} \rangle + (t - t_1) \langle \boldsymbol{m}_h \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle + \langle \widehat{\boldsymbol{m}}_h \lambda_h^1 - \widehat{\boldsymbol{m}}_h^{1/2} \lambda_h^{1/2}, \boldsymbol{\varphi} \rangle \\
&\qquad = \langle (\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h) \lambda_h^1, \boldsymbol{\varphi} \rangle + (t - t_1) \langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle + (t - t_1)^2 \langle \overline{\partial} \boldsymbol{m}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle \\
&\qquad\quad + (t - t_1) \langle \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1 + \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \lambda_h^1, \boldsymbol{\varphi} \rangle + \langle \widehat{\boldsymbol{m}}_h^1 \lambda_h^1 - \widehat{\boldsymbol{m}}_h^{1/2} \lambda_h^{1/2}, \boldsymbol{\varphi} \rangle.
\end{aligned}$$

Moreover, by (2.5.7), adding and subtracting first $\langle \widehat{\boldsymbol{m}}_h^1 \lambda_h^{1/2}, \boldsymbol{\varphi} \rangle$ and second $\frac{\tau}{2} \langle \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \lambda_h^1, \boldsymbol{\varphi} \rangle$, we obtain

$$\begin{aligned}
\langle \widehat{\boldsymbol{m}}_h^1 \lambda_h^1 - \widehat{\boldsymbol{m}}_h^{1/2} \lambda_h^{1/2}, \boldsymbol{\varphi} \rangle &= \frac{\tau}{2} \langle \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle + \frac{\tau}{2} \langle \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \lambda_h^{1/2}, \boldsymbol{\varphi} \rangle \\
&= \frac{\tau}{2} \langle \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1 + \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \lambda_h^1, \boldsymbol{\varphi} \rangle - \frac{\tau^2}{4} \langle \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle.
\end{aligned}$$

Combining the above equations and using $t - t_1 + \frac{\tau}{2} = t - t_{1/2}$ yields

$$\begin{aligned}
\langle \boldsymbol{m}_h \lambda_h - \widehat{\boldsymbol{m}}_h^{1/2} \lambda_h^{1/2}, \boldsymbol{\varphi} \rangle &= \langle (\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h) \lambda_h^1, \boldsymbol{\varphi} \rangle + (t - t_1) \langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle - \frac{\tau^2}{4} \langle \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle \\
&\quad + (t - t_{1/2}) \langle \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1 + \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \lambda_h^1, \boldsymbol{\varphi} \rangle + (t - t_1)^2 \langle \overline{\partial} \boldsymbol{m}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle
\end{aligned} \tag{2.5.10}$$

All in all, we conclude by inserting (2.5.9) and (2.5.10) into the right-hand side of (2.5.8)

$$\begin{aligned}
&\alpha \langle \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi} \rangle + \langle \boldsymbol{m} \times \partial_t \boldsymbol{\mathcal{E}}, \boldsymbol{\varphi} \rangle + \langle \boldsymbol{\epsilon} \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle + a(\widetilde{\boldsymbol{\epsilon}}, \boldsymbol{\varphi}) + b^{\boldsymbol{m}}(\varepsilon, \boldsymbol{\varphi}) + b^{\boldsymbol{\epsilon}}(\lambda_h, \boldsymbol{\varphi}) \\
&\quad = \langle \boldsymbol{f}_h - \boldsymbol{f}, \boldsymbol{\varphi} \rangle + \langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle + (t - t_1) \langle (\overline{\partial} \boldsymbol{m}_h^1 - \overline{\partial} \widehat{\boldsymbol{m}}_h^1) \times \overline{\partial} \boldsymbol{m}_h^1, \boldsymbol{\varphi} \rangle \\
&\qquad + (t - t_1)(t - t_{1/2}) \langle \overline{\partial} \boldsymbol{m}_h^1 \times \overline{\partial}^2 \boldsymbol{m}_h^2, \boldsymbol{\varphi} \rangle + \langle (\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h) \lambda_h^1, \boldsymbol{\varphi} \rangle \\
&\qquad + (t - t_1) \langle (\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle - \frac{\tau^2}{4} \langle \overline{\partial} \widehat{\boldsymbol{m}}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle + (t - t_1)^2 \langle \overline{\partial} \boldsymbol{m}_h^1 \overline{\partial} \lambda_h^1, \boldsymbol{\varphi} \rangle \\
&\qquad + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial} \boldsymbol{m}_h^1), \boldsymbol{\varphi} \rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2} \lambda_h^{1/2}), \boldsymbol{\varphi} \rangle \\
&\qquad + (t - t_{1/2}) \langle \boldsymbol{\Psi}_0, \boldsymbol{\varphi} \rangle \\
&\quad = \langle \boldsymbol{r}_h^0, \boldsymbol{\varphi} \rangle
\end{aligned} \tag{2.5.11}$$

with $\boldsymbol{\Psi}_0 = \alpha\overline{\partial}^2\boldsymbol{m}_h^2 + \overline{\partial}\widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}\boldsymbol{m}_h^1 + \widehat{\boldsymbol{m}}_h^1 \times \overline{\partial}^2\boldsymbol{m}_h^2 - \overline{\partial}(\Delta_h^1\boldsymbol{m}_h^1) + \widehat{\boldsymbol{m}}_h^1\overline{\partial}\lambda_h^1 + \overline{\partial}\widehat{\boldsymbol{m}}_h^1\lambda_h^1 - \overline{\partial}\boldsymbol{f}_h^1$. Following the proof of Theorem 19 and estimating

$$\|\boldsymbol{\nabla}(\boldsymbol{W}-\boldsymbol{m})(0)\| \leq \|\boldsymbol{\nabla}(\boldsymbol{W}-\boldsymbol{M}_h)(0)\| + \|\boldsymbol{\nabla}(\boldsymbol{P}_0^0\boldsymbol{m}(0)-\boldsymbol{m}(0))\|$$

concludes the proof. □

In the following Lemma, we estimate the difference between the linear reconstruction $\boldsymbol{m}_h$ and the three-point reconstruction $\boldsymbol{M}_h$.

Lemma 22. *Let $\boldsymbol{m}_h^n$ be the discrete solution of* (2.3.2) *and $\boldsymbol{M}_h$ be the corresponding three-point reconstruction defined in* (2.1.5) - (2.1.7)*. Then, for $t \in [0, t_N]$ it holds*

$$\|\boldsymbol{\nabla}(\boldsymbol{M}_h-\boldsymbol{m}_h)(t)\|^2 \lesssim \tau^4 \max_{2\leq n\leq N}\|\boldsymbol{\nabla}\overline{\partial}_n^2\boldsymbol{m}_h^n\|^2 + \tau^4 \max_{2\leq n\leq N}\|\boldsymbol{\nabla}(\overline{\partial}^2-\overline{\partial}_n^2)\boldsymbol{m}_h^n\|^2 = \tau^4\mathcal{E}_1 + \Xi_2,$$

where the estimators are defined in Definition 18.

Proof. By the definition of the reconstruction (2.1.5) and (2.1.7), we have

$$\begin{aligned}\boldsymbol{M}_h-\boldsymbol{m}_h &= \frac{1}{2}(t-t_n)(t-t_{n-1})\overline{\partial}^2\boldsymbol{m}_h^n, \quad t\in I_n, n\geq 2,\\ \boldsymbol{M}_h-\boldsymbol{m}_h &= \frac{1}{2}(t-t_0)(t-t_1)\overline{\partial}^2\boldsymbol{m}_h^2, \qquad t\in I_1.\end{aligned} \tag{2.5.12}$$

Thus, by the triangle inequality and (2.5.12), we obtain

$$\|\boldsymbol{\nabla}(\boldsymbol{M}_h-\boldsymbol{m}_h)(t)\|^2 \lesssim \tau^4\|\boldsymbol{\nabla}\overline{\partial}^2\boldsymbol{m}_h^n\|^2 \leq \tau^4\|\boldsymbol{\nabla}\overline{\partial}_n^2\boldsymbol{m}_h^n\|^2 + \tau^4\|\boldsymbol{\nabla}(\overline{\partial}^2-\overline{\partial}_n^2)\boldsymbol{m}_h^n\|^2$$

for $t\in I_n$, $n\geq 2$ and $\|\boldsymbol{\nabla}(\boldsymbol{M}_h-\boldsymbol{m}_h)(t)\|^2 \lesssim \tau^4\|\boldsymbol{\nabla}\overline{\partial}^2\boldsymbol{m}_h^2\|^2$ for $t\in I_1$. □

Next, we derive error bounds for the error between the time-space reconstruction $\boldsymbol{W}$ and the temporal reconstructions, which result in the space error estimators.

Lemma 23. *Let $\boldsymbol{m}_h^n$ be the solution of* (2.3.2)*, $\boldsymbol{M}_h$ be the corresponding three-point reconstructions defined in* (2.1.5) - (2.1.7) *and $\boldsymbol{W}$ be the three-point time-space reconstruction defined in* (2.4.6)*. Then, for $t\in[0,t_N]$ it holds*

$$\|\boldsymbol{\nabla}(\boldsymbol{W}-\boldsymbol{M}_h)(t)\|^2 \lesssim \Lambda_1, \tag{2.5.13}$$

$$\int_0^{t_N}\|\partial_t(\boldsymbol{W}-\boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1}^2\,\mathrm{d}t \lesssim \Lambda_2+\Xi_1, \tag{2.5.14}$$

$$\int_0^{t_N}\|(\boldsymbol{W}-\boldsymbol{m}_h)\|^2\,\mathrm{d}t \lesssim \Lambda_3+\mathcal{E}_6, \tag{2.5.15}$$

where the estimators are defined in Definition 18.

Proof. Recall that $\partial_t\boldsymbol{M}_h(t) = \overline{\partial}^B\boldsymbol{m}_h^n + (t-t_n)\overline{\partial}^2\boldsymbol{m}_h^n$ and $\partial_t\boldsymbol{W}(t) = \overline{\partial}^B\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n + (t-t_n)\overline{\partial}^2\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n$ for $t\in I_n$ and $n\geq 2$, as well as $\overline{\partial}^B(\cdot) = \frac{\tau}{2}\overline{\partial}^2(\cdot) + \overline{\partial}(\cdot)$. Using the definition of the three-point

reconstruction (2.1.5) for $t \in I_n$, $n \geq 2$, along with the discrete differentials given in (2.1.3) and (2.1.6), we obtain

$$\begin{aligned}\partial_t(\boldsymbol{W} - \boldsymbol{M}_h) &= \overline{\partial}^B(\boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n) + (t - t_n)\overline{\partial}^2(\boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n) \\ &= \left(t - t_n + \frac{\tau}{2}\right)\overline{\partial}^2(\boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n) + \overline{\partial}(\boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n) \\ &= \left(\frac{t - t_n}{\tau} + \frac{3}{2}\right)\overline{\partial}(\boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n) - \left(\frac{t - t_n}{\tau} + \frac{1}{2}\right)\overline{\partial}(\boldsymbol{m}_h^{n-1} - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^{n-1}).\end{aligned} \tag{2.5.16}$$

Similarly, applying the definition of the three-point reconstruction (2.1.7) for $t \in I_1$ and (2.1.3) yields

$$\begin{aligned}\partial_t(\boldsymbol{W} - \boldsymbol{M}_h) &= \overline{\partial}(\boldsymbol{m}_h^1 - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^1) + (t - t_{1/2})\overline{\partial}^2(\boldsymbol{m}_h^2 - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^2) \\ &= \frac{t - t_{1/2}}{\tau}\,\overline{\partial}(\boldsymbol{m}_h^2 - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^2) + \left(1 - \frac{t - t_{1/2}}{\tau}\right)\overline{\partial}(\boldsymbol{m}_h^1 - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^1).\end{aligned} \tag{2.5.17}$$

Since $|t - t_n| \leq \tau$ for $t \in I_n$, $n \geq 2$, and $|t - t_{1/2}| \leq \tau$ for $t \in I_1$, we obtain with (2.5.16) and (2.5.17)

$$\int_0^T \|\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1}^2 \,\mathrm{d}t \lesssim \sum_{n=1}^N \tau \|\overline{\partial}(\boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n)\|_{\boldsymbol{\mathcal{H}}^1}^2.$$

Finally, adding and subtracting $\boldsymbol{\mathcal{R}}\overline{\partial}_n \boldsymbol{m}_h^n$ inside $\|\cdot\|_{\boldsymbol{\mathcal{H}}^1}$, using $\overline{\partial}\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n = \boldsymbol{\mathcal{R}}\overline{\partial}\boldsymbol{m}_h^n$ and the triangle inequality leads to

$$\int_0^T \|\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1}^2 \,\mathrm{d}t \lesssim \sum_{n=1}^N \tau \|\overline{\partial}_n \boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\overline{\partial}_n \boldsymbol{m}_h^n\|_{\boldsymbol{\mathcal{H}}^1}^2 + \tau \|(\boldsymbol{I} - \boldsymbol{\mathcal{R}})(\overline{\partial} - \overline{\partial}_n)\boldsymbol{m}_h^n\|_{\boldsymbol{\mathcal{H}}^1}^2. \tag{2.5.18}$$

Estimating $\|\overline{\partial}_n \boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\overline{\partial}_n \boldsymbol{m}_h^n\|_{\boldsymbol{\mathcal{H}}^1}^2 \lesssim \eta(\overline{\partial}_n \boldsymbol{m}_h^n; \mathcal{H}^1)$ via (2.2.6) concludes (2.5.14).

Let $\star \in \{\boldsymbol{\mathcal{L}}^2, \boldsymbol{\mathcal{H}}^1\}$ and $\|\cdot\|_\star$ denote the norm under consideration. Using the definition of the reconstructions, we have

$$\|(\boldsymbol{W} - \boldsymbol{M}_h)(t)\|_\star \lesssim \|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n - \boldsymbol{m}_h^n\|_\star + \|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^{n-1} - \boldsymbol{m}_h^{n-1}\|_\star + \|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^{n-2} - \boldsymbol{m}_h^{n-2}\|_\star \tag{2.5.19}$$

for $t \in I_n$, $n \geq 2$, and

$$\|(\boldsymbol{W} - \boldsymbol{M}_h)(t)\|_\star \lesssim \|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^2 - \boldsymbol{m}_h^2\|_\star + \|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^1 - \boldsymbol{m}_h^1\|_\star + \|\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^0 - \boldsymbol{m}_h^0\|_\star$$

for $t \in I_1$. Combining the above equations with (2.5.12) and (2.2.6), we arrive at

$$\begin{aligned}\int_0^{t_N} \|(\boldsymbol{W} - \boldsymbol{m}_h)\|^2 \,\mathrm{d}t &\lesssim \int_0^{t_N} (\|\boldsymbol{W} - \boldsymbol{M}_h\|^2 + \|\boldsymbol{M}_h - \boldsymbol{m}_h\|^2) \,\mathrm{d}t \\ &\lesssim \sum_{n=0}^N \tau \eta(\boldsymbol{m}_h^n; \mathcal{L}^2) + \sum_{n=2}^N \tau^5 \|\overline{\partial}^2 \boldsymbol{m}_h^n\|^2.\end{aligned}$$

This establishes (2.5.15). Finally, using (2.5.19) and estimating $\|\boldsymbol{\nabla}(\boldsymbol{m}_h^n - \boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n)\|^2 \lesssim \eta(\boldsymbol{m}_h^n; \mathcal{H}^1)$ with (2.2.6) concludes (2.5.13). □

Below, we bound the difference between the linear and quadratic time-space reconstruction similar to Lemma 22.

Lemma 24. *Let $\boldsymbol{m}_h^n$ be the solution of* (2.3.2) *and $\boldsymbol{w}$, $\boldsymbol{W}$ be the corresponding time-space reconstructions defined in* (2.4.5) - (2.4.6)*. Then*

$$\int_0^{t_N} \|\Delta_h(\boldsymbol{w} - \boldsymbol{W})\|^2 \,\mathrm{d}t \lesssim \sum_{n=2}^{N} \tau^5 \|\overline{\partial}^2(-\Delta_h^n)\boldsymbol{m}_h^n\|^2 = \tau^4 \mathcal{E}_5$$

where the estimator is defined in Definition 18.

Proof. Recall that

$$\begin{aligned} -\Delta_h(\boldsymbol{w} - \boldsymbol{W}) &= \frac{1}{2}(t - t_n)(t - t_{n-1})\overline{\partial}^2(-\Delta_h^n)\boldsymbol{m}_h^n, \quad t \in I_n, n \geq 2, \\ -\Delta_h(\boldsymbol{w} - \boldsymbol{W}) &= \frac{1}{2}(t - t_0)(t - t_1)\overline{\partial}^2(-\Delta_h^2)\boldsymbol{m}_h^2, \qquad t \in I_1. \end{aligned}$$

Thus

$$\int_0^{T} \|\Delta_h(\boldsymbol{w} - \boldsymbol{W})\|^2 \,\mathrm{d}t \lesssim \sum_{n=2}^{N} \tau^5 \|\overline{\partial}^2(-\Delta_h^n)\boldsymbol{m}_h^n\|^2.$$

□

Next, we compute bounds for the residuals of the parabolic error equations (2.4.9) and (2.5.11).

Lemma 25. *Let $\boldsymbol{r}_h$ be defined by* (2.4.10) *and $\boldsymbol{r}_h^0$ by* (2.5.1)*. Then,*

$$\begin{aligned} \int_0^{t_1} \|\boldsymbol{r}_h^0\|^2 \,\mathrm{d}t &\lesssim \mathcal{F}_1 + \mathcal{C}_1 + \mathcal{Q}_1 + \mathcal{I}_2 + \mathcal{I}_3, \\ \int_{t_1}^{t_N} \|\boldsymbol{r}_h\|^2 \,\mathrm{d}t &\lesssim \mathcal{F}_2 + \tau^4(\mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4) + \Xi_3 + \mathcal{C}_2 + \mathcal{Q}_2 + \mathcal{I}_1, \end{aligned}$$

where the estimators are defined in Definition 18.

Proof. Recall that

$$\begin{aligned} \langle \boldsymbol{r}_h, \boldsymbol{\varphi} \rangle = {} & \langle \boldsymbol{f}_h^n - \boldsymbol{f}, \boldsymbol{\varphi} \rangle + \alpha \langle (\overline{\partial}^B - \overline{\partial}_n^B)\boldsymbol{m}_h^n, \boldsymbol{\varphi} \rangle + \langle (\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n) \times \partial_t \boldsymbol{M}_h, \boldsymbol{\varphi} \rangle \\ & + (t - t_n)\langle (\overline{\partial}\boldsymbol{m}_h^n - \overline{\partial}\widehat{\boldsymbol{m}}_h^n) \times \overline{\partial}^B \boldsymbol{m}_h^n, \boldsymbol{\varphi} \rangle + (t - t_n)^2 \langle \overline{\partial}\boldsymbol{m}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n, \boldsymbol{\varphi} \rangle \\ & + \langle (\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^n, \boldsymbol{\varphi} \rangle + (t - t_n)\langle (\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n)\overline{\partial}\lambda_h^n, \boldsymbol{\varphi} \rangle \\ & + (t - t_n)^2 \langle \overline{\partial}\boldsymbol{m}_h^n \overline{\partial}\lambda_h^n, \boldsymbol{\varphi} \rangle + (t - t_n)\langle \boldsymbol{\Psi}, \boldsymbol{\varphi} \rangle \\ & + \langle (\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n), \boldsymbol{\varphi} \rangle + \langle (\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n), \boldsymbol{\varphi} \rangle, \end{aligned}$$

for all $\boldsymbol{\varphi} \in \mathcal{L}^2(0,t_1;\boldsymbol{V})$, where $\boldsymbol{\Psi}$ is given by (2.4.11), for $t \in I_n$ and $n \geq 2$. Although $\boldsymbol{\Psi}$ is already computable a posteriori, we further simplify it to clarify different error components and obtain more manageable terms. We start by using equation (2.4.12), i.e.,

$$\alpha\langle \overline{\partial}_n^B \boldsymbol{m}_h^n + \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n) - \Delta_h^n \boldsymbol{m}_h^n + \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \lambda_h^n) - \boldsymbol{f}_h^n, \boldsymbol{\varphi}\rangle = 0, \qquad \forall \boldsymbol{\varphi} \in \mathcal{L}^2(0,t_1;\boldsymbol{V}),$$

to obtain

$$\begin{aligned}
\langle \boldsymbol{\Psi}, \boldsymbol{\varphi}\rangle &= \langle \alpha\overline{\partial}^2 \boldsymbol{m}_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \overline{\partial}\lambda_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n \lambda_h^n, \boldsymbol{\varphi}\rangle \\
&\quad + \frac{1}{\tau}\langle -\alpha\overline{\partial}_n^B \boldsymbol{m}_h^n - \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n) - \boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \lambda_h^n) + \boldsymbol{f}_h^n, \boldsymbol{\varphi}\rangle \\
&\quad + \frac{1}{\tau}\langle \alpha\overline{\partial}_{n-1}^B \boldsymbol{m}_h^{n-1} + \boldsymbol{P}_0^{n-1}(\widehat{\boldsymbol{m}}_h^{n-1} \times \overline{\partial}^B \boldsymbol{m}_h^{n-1}) + \boldsymbol{P}_0^{n-1}(\widehat{\boldsymbol{m}}_h^{n-1} \lambda_h^{n-1}) - \boldsymbol{f}_h^{n-1}, \boldsymbol{\varphi}\rangle \\
&= \langle \overline{\partial}\boldsymbol{f}_h^n, \boldsymbol{\varphi}\rangle + \alpha\langle \overline{\partial}^2 \boldsymbol{m}_h^n - \frac{1}{\tau}(\overline{\partial}_n^B \boldsymbol{m}_h^n - \overline{\partial}_{n-1}^B \boldsymbol{m}_h^{n-1}), \boldsymbol{\varphi}\rangle + \langle \widehat{\boldsymbol{m}}_h^n \overline{\partial}\lambda_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n \lambda_h^n, \boldsymbol{\varphi}\rangle \\
&\quad - \langle \overline{\partial}(\boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \lambda_h^n)), \boldsymbol{\varphi}\rangle + \langle \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n - \overline{\partial}(\boldsymbol{P}_0^n(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n)), \boldsymbol{\varphi}\rangle
\end{aligned}$$

for $t \in I_n$, $n \geq 3$. Since $\overline{\partial}_n^B \boldsymbol{m}_h^n = \frac{\tau}{2}\overline{\partial}_n^2 \boldsymbol{m}_h^n + \overline{\partial}_n \boldsymbol{m}_h^n$ by (2.1.6), we have

$$\begin{aligned}
\overline{\partial}^2 \boldsymbol{m}_h^n - \frac{1}{\tau}(\overline{\partial}_n^B \boldsymbol{m}_h^n - \overline{\partial}_{n-1}^B \boldsymbol{m}_h^{n-1}) &= \overline{\partial}^2 \boldsymbol{m}_h^n - \frac{\overline{\partial}_n \boldsymbol{m}_h^n - \overline{\partial}_{n-1}\boldsymbol{m}_h^{n-1}}{\tau} - \frac{1}{2}(\overline{\partial}_n^2 \boldsymbol{m}_h^n - \overline{\partial}_{n-1}^2 \boldsymbol{m}_h^{n-1}) \\
&= \overline{\partial}(\overline{\partial} - \overline{\partial}_n)\boldsymbol{m}_h^n - \frac{1}{2}(\overline{\partial}_n^2 \boldsymbol{m}_h^n - \overline{\partial}_{n-1}^2 \boldsymbol{m}_h^{n-1}).
\end{aligned} \tag{2.5.20}$$

Applying (2.5.20) and adding and subtracting $\langle \overline{\partial}(\widehat{\boldsymbol{m}}_h^n \lambda_h^n) + \overline{\partial}(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n), \boldsymbol{\varphi}\rangle$ leads to

$$\begin{aligned}
\langle \boldsymbol{\Psi}, \boldsymbol{\varphi}\rangle &= \langle \overline{\partial}\boldsymbol{f}_h^n, \boldsymbol{\varphi}\rangle - \frac{\alpha}{2}\langle \overline{\partial}_n^2 \boldsymbol{m}_h^n - \overline{\partial}_{n-1}^2 \boldsymbol{m}_h^{n-1}, \boldsymbol{\varphi}\rangle + \alpha\langle \overline{\partial}((\overline{\partial} - \overline{\partial}_n)\boldsymbol{m}_h^n), \boldsymbol{\varphi}\rangle \\
&\quad + \langle \overline{\partial}((\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n)), \boldsymbol{\varphi}\rangle + \langle \widehat{\boldsymbol{m}}_h^n \overline{\partial}\lambda_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n \lambda_h^n - \overline{\partial}(\widehat{\boldsymbol{m}}_h^n \lambda_h^n), \boldsymbol{\varphi}\rangle \\
&\quad + \langle \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n - \overline{\partial}(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n), \boldsymbol{\varphi}\rangle + \langle \overline{\partial}((\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n)), \boldsymbol{\varphi}\rangle.
\end{aligned} \tag{2.5.21}$$

Using that

$$\overline{\partial}(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n) = \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}(\overline{\partial}^B \boldsymbol{m}_h^n) - \tau\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}(\overline{\partial}^B \boldsymbol{m}_h^n)$$

yields

$$\begin{aligned}
&\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n - \overline{\partial}(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n) \\
&\quad = \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n + \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n \\
&\qquad - \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}(\overline{\partial}^B \boldsymbol{m}_h^n) + \tau\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}(\overline{\partial}^B \boldsymbol{m}_h^n) \\
&\quad = -\frac{\tau}{2}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^3 \boldsymbol{m}_h^n + \frac{\tau^2}{2}\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^3 \boldsymbol{m}_h^n + \tau\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n \\
&\quad = -\frac{\tau}{2}\widehat{\boldsymbol{m}}_h^{n-1} \times \overline{\partial}^3 \boldsymbol{m}_h^n + \tau\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n.
\end{aligned} \tag{2.5.22}$$

Moreover, we have

$$\overline{\partial}(\widehat{\boldsymbol{m}}_h^n\lambda_h^n) = \overline{\partial}\widehat{\boldsymbol{m}}_h^n\lambda_h^n + \widehat{\boldsymbol{m}}_h^{n-1}\overline{\partial}\lambda_h^n = \overline{\partial}\widehat{\boldsymbol{m}}_h^n\lambda_h^n + \widehat{\boldsymbol{m}}_h^n\overline{\partial}\lambda_h^n - \tau\overline{\partial}\widehat{\boldsymbol{m}}_h^n\overline{\partial}\lambda_h^n.$$

Combining the above equations with (2.5.21), we conclude

$$\begin{aligned}\langle\boldsymbol{\Psi},\boldsymbol{\varphi}\rangle &= \langle\overline{\partial}\boldsymbol{f}_h^n,\boldsymbol{\varphi}\rangle - \frac{\alpha}{2}\langle\overline{\partial}_n^2\boldsymbol{m}_h^n - \overline{\partial}_{n-1}^2\boldsymbol{m}_h^{n-1},\boldsymbol{\varphi}\rangle + \alpha\langle\overline{\partial}((\overline{\partial}-\overline{\partial}_n)\boldsymbol{m}_h^n),\boldsymbol{\varphi}\rangle + \tau\langle\overline{\partial}\widehat{\boldsymbol{m}}_h^n\overline{\partial}\lambda_h^n,\boldsymbol{\varphi}\rangle\\ &\quad + \langle\overline{\partial}((\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n)),\boldsymbol{\varphi}\rangle + \langle\overline{\partial}((\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\lambda_h^n)),\boldsymbol{\varphi}\rangle\\ &\quad - \frac{\tau}{2}\langle\widehat{\boldsymbol{m}}_h^{n-1}\times\overline{\partial}^3\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle + \tau\langle\overline{\partial}\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n,\boldsymbol{\varphi}\rangle =: g^n(t).\end{aligned}$$

For the case $t\in I_2$, we have due to the definition of $\boldsymbol{m}_h^{-1}$ given by (2.4.1) and the trapezoidal scheme (2.3.7)

$$\begin{aligned}\alpha\overline{\partial}^B\boldsymbol{m}_h^1 &= \frac{\alpha}{\tau}(\frac{3}{2}\boldsymbol{m}_h^1 - 2\boldsymbol{m}_h^0 + \frac{1}{2}\boldsymbol{m}_h^{-1})\\ &= 2\alpha\overline{\partial}\boldsymbol{m}_h^1 - \boldsymbol{f}_h^0 - \Delta_h^0\boldsymbol{m}_h^0 + \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^0\times\overline{\partial}\boldsymbol{m}_h^1) + \boldsymbol{P}_0^1(\lambda_h^0\widehat{\boldsymbol{m}}_h^0)\\ &= \boldsymbol{f}_h^1 + \Delta_h^1\boldsymbol{m}_h^1 - \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^1\times\overline{\partial}\boldsymbol{m}_h^1) + \boldsymbol{P}_0^1(\lambda_h^0\widehat{\boldsymbol{m}}_h^0) - 2\boldsymbol{P}_0^1(\lambda_h^{1/2}\widehat{\boldsymbol{m}}_h^{1/2}).\end{aligned}\tag{2.5.23}$$

Using that $\lambda_h^0\widehat{\boldsymbol{m}}_h^0 + \lambda_h^1\widehat{\boldsymbol{m}}_h^1 - 2\lambda_h^{1/2}\widehat{\boldsymbol{m}}_h^{1/2} = \frac{1}{2}\tau^2\overline{\partial}\lambda_h^1\overline{\partial}\widehat{\boldsymbol{m}}_h^1$, we get

$$\begin{aligned}\alpha\overline{\partial}^B\boldsymbol{m}_h^1 &= \boldsymbol{f}_h^1 + \Delta_h^1\boldsymbol{m}_h^1 - \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^1\times\overline{\partial}\boldsymbol{m}_h^1) - \boldsymbol{P}_0^1(\lambda_h^1\widehat{\boldsymbol{m}}_h^1) + \frac{1}{2}\tau^2\boldsymbol{P}_0^1(\overline{\partial}\lambda_h^1\overline{\partial}\widehat{\boldsymbol{m}}_h^1)\\ &= \boldsymbol{f}_h^1 + \Delta_h^1\boldsymbol{m}_h^1 - \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^1\times\overline{\partial}^B\boldsymbol{m}_h^1) - \boldsymbol{P}_0^1(\lambda_h^1\widehat{\boldsymbol{m}}_h^1) + \frac{1}{2}\tau^2\boldsymbol{P}_0^1(\overline{\partial}\lambda_h^1\overline{\partial}\widehat{\boldsymbol{m}}_h^1)\\ &\quad + \frac{\tau}{2}\boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^1\times\overline{\partial}^2\boldsymbol{m}_h^1),\end{aligned}\tag{2.5.24}$$

where $\alpha\overline{\partial}^2\boldsymbol{m}_h^1 = \frac{\alpha}{\tau^2}(\boldsymbol{m}_h^1 - 2\boldsymbol{m}_h^0 + \boldsymbol{m}_h^{-1})$ with $\boldsymbol{m}_h^{-1}$ given by (2.4.1). Thus, we obtain for $t\in I_2$ using (2.5.24)

$$\langle\boldsymbol{\Psi},\boldsymbol{\varphi}\rangle = g^2(t) + \frac{\tau}{2}\langle\boldsymbol{P}_0^1(\overline{\partial}\lambda_h^1\overline{\partial}\widehat{\boldsymbol{m}}_h^1)\rangle + \frac{1}{2}\langle\boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^1\times\overline{\partial}^2\boldsymbol{m}_h^1),\boldsymbol{\varphi}\rangle.$$

Finally, the triangle inequality yields for the residual $\boldsymbol{r}_h$ defined in (2.4.10)

$$\begin{aligned}\|\boldsymbol{r}_h\| \lesssim\ & \|\boldsymbol{f}_h - \boldsymbol{f}\| + \|(\overline{\partial}^B - \overline{\partial}_n^B)\boldsymbol{m}_h^n\| + \|(\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n)\times\partial_t\boldsymbol{M}_h\|\\ &+ \tau\|(\overline{\partial}\boldsymbol{m}_h^n - \overline{\partial}\widehat{\boldsymbol{m}}_h^n)\times\overline{\partial}^B\boldsymbol{m}_h^n\| + \tau^2\|\overline{\partial}\boldsymbol{m}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n\|\\ &+ \|(\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^n\| + \tau\|(\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n)\overline{\partial}\lambda_h^n\| + \tau^2\|\overline{\partial}\boldsymbol{m}_h^n\overline{\partial}\lambda_h^n\|\\ &+ \|(\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n)\| + \|(\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\lambda_h^n)\|\\ &+ \tau\|\overline{\partial}_n^2\boldsymbol{m}_h^n - \overline{\partial}_{n-1}^2\boldsymbol{m}_h^{n-1}\| + \tau\|\overline{\partial}((\overline{\partial}-\overline{\partial}_n)\boldsymbol{m}_h^n)\| + \tau^2\|\overline{\partial}\widehat{\boldsymbol{m}}_h^n\overline{\partial}\lambda_h^n\|\\ &+ \tau\|\overline{\partial}((\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^B\boldsymbol{m}_h^n))\| + \tau\|\overline{\partial}((\boldsymbol{I}-\boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n\lambda_h^n))\|\\ &+ \tau^2\|\widehat{\boldsymbol{m}}_h^{n-1}\times\overline{\partial}^3\boldsymbol{m}_h^n\| + \tau^2\|\overline{\partial}\widehat{\boldsymbol{m}}_h^n\times\overline{\partial}^2\boldsymbol{m}_h^n\|\\ &+ \tau^2\chi_{I_2}(t)\|\boldsymbol{P}_0^1(\overline{\partial}\lambda_h^1\overline{\partial}\widehat{\boldsymbol{m}}_h^1)\| + \tau\chi_{I_2}(t)\|\boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^1\times\overline{\partial}^2\boldsymbol{m}_h^1)\|,\end{aligned}$$

for $t \in I_n$ and $n \geq 2$, where χ_{I_2} denotes the indicator function onto I_2. Integrating from $t = t_{k-1}$ to t_k and summing over $k = 2$ to N concludes the proof for $\boldsymbol{r}_h$.

Using the definition of $\boldsymbol{r}_h^0$ (2.5.1), we obtain

$$\begin{aligned}\|\boldsymbol{r}_h^0\| \lesssim{}& \|\boldsymbol{f}_h - \boldsymbol{f}\| + \|(\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1) \times \partial_t \boldsymbol{M}_h\| + \tau\|(\overline{\partial}\boldsymbol{m}_h^1 - \overline{\partial}\widehat{\boldsymbol{m}}_h^1) \times \overline{\partial}\boldsymbol{m}_h^1\| \\ &+ \tau^2\|\overline{\partial}\boldsymbol{m}_h^1 \times \overline{\partial}^2\boldsymbol{m}_h^2\| + \|(\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^1\| + \tau\|(\boldsymbol{m}_h^1 - \widehat{\boldsymbol{m}}_h^1)\overline{\partial}\lambda_h^1\| + \tau^2\|\overline{\partial}\widehat{\boldsymbol{m}}_h^1\overline{\partial}\lambda_h^1\| \\ &+ \tau^2\|\overline{\partial}\boldsymbol{m}_h^1\overline{\partial}\lambda_h^1\| + \|(\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2} \times \overline{\partial}\boldsymbol{m}_h^1)\| + \|(\boldsymbol{I} - \boldsymbol{P}_0^1)(\widehat{\boldsymbol{m}}_h^{1/2}\lambda_h^{1/2})\| + \tau\|\boldsymbol{\Psi}_0\|.\end{aligned}$$

□

We proceed with the $\mathcal{L}^\infty$-error bounds for the error between the time-space and temporal three-point reconstruction.

Lemma 26. *Let $\boldsymbol{m}_h^n$ be the solution of* (2.3.2)*, $\boldsymbol{M}_h$ be the corresponding three-point reconstructions defined in* (2.1.5) - (2.1.7) *and $\boldsymbol{W}$ be the three-point time-space reconstruction defined in* (2.4.6)*. Then*

$$\|\boldsymbol{W} - \boldsymbol{M}_h\|^2_{\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{L}}^\infty)} \lesssim \max_{0\leq n\leq N} \eta(\boldsymbol{m}_h^n; \mathcal{L}^\infty), \tag{2.5.25}$$

$$\|\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|^2_{\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{L}}^\infty)} \lesssim \max_{1\leq n\leq N} \eta(\overline{\partial}_n\boldsymbol{m}_h^n; \mathcal{L}^\infty) + \max_{1\leq n\leq N} \eta((\overline{\partial} - \overline{\partial}_n)\boldsymbol{m}_h^n; \mathcal{L}^\infty), \tag{2.5.26}$$

$$\|\boldsymbol{\nabla}(\boldsymbol{W} - \boldsymbol{M}_h)\|^2_{\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{L}}^\infty)} \lesssim \max_{0\leq n\leq N} \eta(\boldsymbol{m}_h^n; \mathcal{W}^{1,\infty}), \tag{2.5.27}$$

$$\|\boldsymbol{\nabla}\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|^2_{\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{L}}^\infty)} \lesssim \max_{1\leq n\leq N} \eta(\overline{\partial}_n\boldsymbol{m}_h^n; \mathcal{W}^{1,\infty}) + \max_{1\leq n\leq N} \eta((\overline{\partial} - \overline{\partial}_n)\boldsymbol{m}_h^n; \mathcal{W}^{1,\infty}), \tag{2.5.28}$$

where we employ the estimators from Assumption 15.

Proof. Combining (2.5.19) and (2.2.8) yields (2.5.25). Following (2.5.18), we obtain (2.5.26). Finally, (2.5.27) and (2.5.28) follow in the same way. □

Finally, we compute the a posteriori bound for the normalization error of the time-space reconstruction.

Lemma 27. *Let $\boldsymbol{m}_h^n$ be the solution of* (2.3.2)*, $\boldsymbol{M}_h$ be the corresponding three-point reconstructions defined in* (2.1.5) - (2.1.7) *and $\boldsymbol{W}$ be the three-point time-space reconstruction defined in* (2.4.6)*. Then*

$$\begin{aligned}&\int_0^{t_N} \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t\boldsymbol{W}\|^2_{\boldsymbol{\mathcal{H}}^1}\,\mathrm{d}t \\ &\quad\lesssim \int_0^{t_N} \left(\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{M}_h))\partial_t\boldsymbol{M}_h\|^2_{\boldsymbol{\mathcal{H}}^1} + \|\boldsymbol{M}_h - \boldsymbol{W}\|^2_{\boldsymbol{\mathcal{H}}^1} + \|\partial_t(\boldsymbol{M}_h - \boldsymbol{W})\|^2_{\boldsymbol{\mathcal{H}}^1}\right)\mathrm{d}t, \\ &\quad\lesssim \mathcal{P} + \Lambda_2 + \Lambda_3 + \Xi_1,\end{aligned}$$

where the (hidden) constant depends on $\|\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}$, $\|\partial_t\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}$, $\eta(\boldsymbol{m}_h^n; \mathcal{W}^{1,\infty})$ and where the estimators are defined in Definition 18.

Proof. By adding and subtracting $(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{M}_h + \boldsymbol{P}(\boldsymbol{M}_h)\partial_t \boldsymbol{M}_h$ and using the triangle inequality, we obtain

$$\begin{aligned}\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{W}\|_{\boldsymbol{\mathcal{H}}^1} &\leq \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{H}}^1} + \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1} \\ &\leq \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{M}_h))\partial_t \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{H}}^1} + \|(\boldsymbol{P}(\boldsymbol{M}_h) - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{H}}^1} \\ &\quad + \|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1}.\end{aligned}$$

Since $\boldsymbol{P}(\boldsymbol{W}) = \boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^\top$, we estimate

$$\|(\boldsymbol{I} - \boldsymbol{P}(\boldsymbol{W}))\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1} = \|\boldsymbol{W}\boldsymbol{W}^\top \partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1} \leq \|\boldsymbol{W}\|^2_{\boldsymbol{\mathcal{W}}^{1,\infty}} \|\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1}.$$

Using the Lipschitz-type bound of the orthogonal projection in Lemma 16, the a posteriori bound of the elliptic reconstruction (2.2.6) and (2.5.19), we obtain

$$\int_0^{t_N} \|(\boldsymbol{P}(\boldsymbol{M}_h) - \boldsymbol{P}(\boldsymbol{W}))\partial_t \boldsymbol{M}_h\|^2_{\boldsymbol{\mathcal{H}}^1} \,\mathrm{d}t \leq \int_0^{t_N} \gamma(\boldsymbol{M}_h, \partial_t \boldsymbol{M}_h)^2 \|\boldsymbol{M}_h - \boldsymbol{W}\|^2_{\boldsymbol{\mathcal{H}}^1} \,\mathrm{d}t \lesssim \Lambda_3.$$

Finally, applying Lemma 23 to estimate $\|\partial_t(\boldsymbol{W} - \boldsymbol{M}_h)\|_{\boldsymbol{\mathcal{H}}^1}$ and bounding $\|\boldsymbol{W}\|^2_{\boldsymbol{\mathcal{W}}^{1,\infty}}$ by the triangle inequality

$$\|\boldsymbol{W}\|_{\boldsymbol{\mathcal{W}}^{1,\infty}} \leq \|\boldsymbol{W} - \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}} + \|\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}, \tag{2.5.29}$$

where $\|\boldsymbol{W} - \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}$ is bounded by (2.5.27), concludes the proof. □

Applying the triangle inequality as in (2.5.29) and Lemma 26, we obtain a posteriori computable error bounds for the coefficients $\gamma_{0,\boldsymbol{W}}$ and $\gamma_{1,\boldsymbol{W}}$ defined in (2.4.24).

2.6 VARIABLE TIME-STEPS

In this section, we extend the a posteriori analysis to the case of variable time-steps. Similar to [25, Sct. 7.1], the proofs may be directly extended after adapting the definitions of the three-point reconstruction and the BDF(2) operator $\overline{\partial}^B(\cdot)$. We therefore present only the resulting changes to the error indicators from Definition 18, after first recalling the precise definitions of the three-point reconstruction and the BDF(2) operator. For simplicity, we assume a fixed spatial mesh. However, the extension to variable spatial meshes is straightforward.

Let $0 = t_0 < t_1 < \cdots < t_N = T$ be a partition of $[0, T]$ with time-steps $\tau_n \coloneqq t_n - t_{n-1}$. Further define the step ratio $\kappa_n \coloneqq \tau_n / \tau_{n-1}$ for $n \geq 1$. The definition of the finite-difference operator $\overline{\partial}$ is extended to

$$\overline{\partial} \boldsymbol{m}_h^n \coloneqq \overline{\partial}^1 \boldsymbol{m}_h^n \coloneqq \frac{\boldsymbol{m}_h^n - \boldsymbol{m}_h^{n-1}}{\tau_n}, \qquad \overline{\partial}^k \boldsymbol{m}_h^n \coloneqq \frac{\overline{\partial}^{k-1} \boldsymbol{m}_h^n - \overline{\partial}^{k-1} \boldsymbol{m}_h^{n-1}}{\tau_n} \tag{2.6.1}$$

for $k \geq 2$. Then, the linear interpolation polynomial is again defined by

$$\boldsymbol{m}_h(t) = \boldsymbol{m}_h^n + (t - t_n)\overline{\partial}\boldsymbol{m}_h^n$$

for $t \in I_n \coloneqq (t_{n-1}, t_n]$ and $n \geq 1$. In contrast to (2.1.5), the three-point reconstruction is modified to

$$\begin{aligned} \boldsymbol{M}_h(t) &\coloneqq \boldsymbol{m}_h(t) + \frac{\kappa_n}{\kappa_n + 1}(t - t_n)(t - t_{n-1})\overline{\partial}^2\boldsymbol{m}_h^n, \\ \partial_t\boldsymbol{M}_h(t) &= \overline{\partial}^B\boldsymbol{m}_h^n + \frac{2\kappa_n}{\kappa_n + 1}(t - t_n)\overline{\partial}^2\boldsymbol{m}_h^n \end{aligned} \tag{2.6.2}$$

for $t \in I_n \coloneqq (t_{n-1}, t_n]$ and $n \geq 2$, while the BDF(2) time discretization operator is defined as

$$\overline{\partial}^B\boldsymbol{m}_h^n \coloneqq \frac{1}{\tau_n}\Big(\frac{1 + 2\kappa_n}{1 + \kappa_n}\boldsymbol{m}_h^n - (1 + \kappa_n)\boldsymbol{m}_h^{n-1} + \frac{\kappa_n^2}{1 + \kappa_n}\boldsymbol{m}_h^{n-2}\Big) = \tau_n\frac{\kappa_n}{1 + \kappa_n}\overline{\partial}^2\boldsymbol{m}_h^n + \overline{\partial}\boldsymbol{m}_h^n. \tag{2.6.3}$$

These definitions are based on [25, Sct. 7.1]. For $t \in I_1$, we again take the three-point reconstruction as defined in (2.1.7). By following the argument in Lemma 20, we derive for variable step-sizes, instead of $\boldsymbol{\Psi}$,

$$\widetilde{\boldsymbol{\Psi}} \coloneqq \alpha\frac{2\kappa_n}{\kappa_n + 1}\overline{\partial}^2\boldsymbol{m}_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B\boldsymbol{m}_h^n + \frac{2\kappa_n}{\kappa_n + 1}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2\boldsymbol{m}_h^n - \overline{\partial}(\Delta_h^n\boldsymbol{m}_h^n) + \widehat{\boldsymbol{m}}_h^n\overline{\partial}\lambda_h^n + \overline{\partial}\widehat{\boldsymbol{m}}_h^n\lambda_h^n.$$

Hence, the error indicators from Lemma 25 need to be adapted. Rewriting (2.5.20) in the case of non-uniform time-steps using (2.6.3), we obtain

$$\begin{aligned} &\frac{2\kappa_n}{\kappa_n + 1}\overline{\partial}^2\boldsymbol{m}_h^n - \frac{1}{\tau_n}\Big(\overline{\partial}^B\boldsymbol{m}_h^n - \overline{\partial}^B\boldsymbol{m}_h^{n-1}\Big) \\ &\qquad = \Big(\frac{2\kappa_n}{\kappa_n + 1} - \frac{\kappa_n}{\kappa_n + 1} - 1\Big)\overline{\partial}^2\boldsymbol{m}_h^n + \frac{1}{\kappa_n}\frac{\kappa_{n-1}}{\kappa_{n-1} + 1}\overline{\partial}^2\boldsymbol{m}_h^{n-1} \\ &\qquad = \frac{1}{\kappa_n}\Big(-\frac{\kappa_n}{\kappa_n + 1}\overline{\partial}^2\boldsymbol{m}_h^n + \frac{\kappa_{n-1}}{\kappa_{n-1} + 1}\overline{\partial}^2\boldsymbol{m}_h^{n-1}\Big). \end{aligned} \tag{2.6.4}$$

Finally, employing the same arguments as in the proof of Lemma 25 and applying (2.6.4), the error indicator $\tau^4\mathcal{E}_2$ from Definition 18 is modified to

$$\widetilde{\mathcal{E}}_2 \coloneqq \sum_{n=2}^{N}\frac{\tau_n^3}{\kappa_n^2}\|\frac{\kappa_n}{\kappa_n + 1}\overline{\partial}^2\boldsymbol{m}_h^n - \frac{\kappa_{n-1}}{\kappa_{n-1} + 1}\overline{\partial}^2\boldsymbol{m}_h^{n-1}\|^2.$$

Moreover, rewriting equation (2.5.22) using (2.6.3) leads to

$$\begin{aligned} &\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B\boldsymbol{m}_h^n + \frac{2\kappa_n}{\kappa_n + 1}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2\boldsymbol{m}_h^n - \overline{\partial}(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B\boldsymbol{m}_h^n) \\ &\quad = \frac{2\kappa_n}{\kappa_n + 1}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}(\overline{\partial}^B\boldsymbol{m}_h^n) + \tau_n\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}(\overline{\partial}^B\boldsymbol{m}_h^n) \\ &\quad = \frac{2\kappa_n}{\kappa_n + 1}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2\boldsymbol{m}_h^n - \tau_n\frac{\kappa_n}{1 + \kappa_n}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^3\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2\boldsymbol{m}_h^n + \tau_n\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}(\overline{\partial}^B\boldsymbol{m}_h^n) \\ &\quad = \frac{\kappa_n - 1}{\kappa_n + 1}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2\boldsymbol{m}_h^n - \tau_n\frac{\kappa_n}{1 + \kappa_n}\widehat{\boldsymbol{m}}_h^{n-1} \times \overline{\partial}^3\boldsymbol{m}_h^n + \tau_n\overline{\partial}\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2\boldsymbol{m}_h^n. \end{aligned}$$

This yields the additional error estimator

$$\widetilde{\mathcal{E}}_7 \coloneqq \sum_{n=2}^{N} \tau_n^3 \frac{(\kappa_n - 1)^2}{(\kappa_n + 1)^2} \|\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n\|^2.$$

The remaining estimators are obtained analogously to the uniform time-step case. For completeness, we summarize the definitions of the error indicators for variable step sizes in the following definition, cf. Definition 18.

Definition 28 (Error estimators for variable time-steps)**.** *Let* $\{\boldsymbol{m}_h^n\}_{n=0}^N$, $\{\widehat{\boldsymbol{m}}_h^n\}_{n=0}^N$, $\{\boldsymbol{f}_h^n\}_{n=0}^N$, $\{\lambda_h^n\}_{n=0}^N$ *be sequences with* $\boldsymbol{m}_h^n, \boldsymbol{f}_h^n \in \boldsymbol{V}_h^n$, $\lambda_h^n \in V_h^n$ *and* $\widehat{\boldsymbol{m}}_h^n \in V$ *for all* $n \geq 0$ *and define*

$$\alpha \boldsymbol{m}_h^{-1} \coloneqq \alpha \boldsymbol{m}_h^1 - 2\tau_1 (\boldsymbol{f}_h^0 + \Delta_h^0 \boldsymbol{m}_h^0 - \boldsymbol{P}_0^1(\widehat{\boldsymbol{m}}_h^0 \times \overline{\partial} \boldsymbol{m}_h^1) - \boldsymbol{P}_0^1(\lambda_h^0 \widehat{\boldsymbol{m}}_h^0)). \tag{2.6.5}$$

With the definition of $\boldsymbol{M}_h$ *in* (2.6.2)*, we introduce the* ***time error estimators***

$$\widetilde{\mathcal{E}}_1 \coloneqq \max_{2 \leq n \leq N} \tau_n^4 \|\boldsymbol{\nabla} \overline{\partial}_n^2 \boldsymbol{m}_h^n\|^2, \qquad \widetilde{\mathcal{E}}_2 \coloneqq \sum_{n=2}^{N} \frac{\tau_n^3}{\kappa_n^2} \|\frac{\kappa_n}{\kappa_n + 1} \overline{\partial}^2 \boldsymbol{m}_h^n - \frac{\kappa_{n-1}}{\kappa_{n-1} + 1} \overline{\partial}^2 \boldsymbol{m}_h^{n-1}\|^2,$$

$$\widetilde{\mathcal{E}}_3 \coloneqq \sum_{n=1}^{N} \tau_n^5 (\|\overline{\partial} \boldsymbol{m}_h^n \overline{\partial} \lambda_h^n\|^2 + \|\overline{\partial} \widehat{\boldsymbol{m}}_h^n \overline{\partial} \lambda_h^n\|^2),$$

$$\widetilde{\mathcal{E}}_4 \coloneqq \sum_{n=2}^{N} \tau_n^5 \Big(\|\overline{\partial} \boldsymbol{m}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n\|^2 + \|\overline{\partial} \widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n\|^2 + \|\widehat{\boldsymbol{m}}_h^{n-1} \times \overline{\partial}^3 \boldsymbol{m}_h^n\|^2 \Big),$$

$$\widetilde{\mathcal{E}}_6 \coloneqq \sum_{n=2}^{N} \tau_n^5 \|\overline{\partial}^2 \boldsymbol{m}_h^n\|^2, \qquad \widetilde{\mathcal{E}}_7 \coloneqq \sum_{n=2}^{N} \tau_n^3 \frac{(\kappa_n - 1)^2}{(\kappa_n + 1)^2} \|\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^2 \boldsymbol{m}_h^n\|^2,$$

the ***reconstruction error estimator***

$$\widetilde{\mathcal{E}}_5 \coloneqq \sum_{n=2}^{N} \tau_n^5 \|\overline{\partial}^2 (-\Delta_h^n) \boldsymbol{m}_h^n\|^2,$$

the ***space error estimators***

$$\widetilde{\Lambda}_2 \coloneqq \sum_{n=1}^{N} \tau_n \, \eta(\overline{\partial}_n \boldsymbol{m}_h^n; \mathcal{H}^1), \qquad \widetilde{\Lambda}_3 \coloneqq \sum_{n=0}^{N} \tau_n \, \eta(\boldsymbol{m}_h^n; \mathcal{L}^2),$$

where η *is the elliptic a posteriori estimator defined in* (2.2.7)*, the* ***finite element space conforming estimators***

$$\begin{aligned} \widetilde{\mathcal{C}}_2 \coloneqq \sum_{n=2}^{N} \Big(& \tau_n \|(\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n)\|^2 + \tau_n \|(\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n)\|^2 \\ & + \tau_n^3 \|\overline{\partial}((\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \lambda_h^n))\|^2 + \tau_n^3 \|\overline{\partial}((\boldsymbol{I} - \boldsymbol{P}_0^n)(\widehat{\boldsymbol{m}}_h^n \times \overline{\partial}^B \boldsymbol{m}_h^n))\|^2 \Big) \end{aligned}$$

and the ***extrapolation error estimators***

$$\widetilde{\mathcal{Q}}_2 \coloneqq \sum_{n=2}^{N} \Big(\int_{I_n} (\|(\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n) \times \partial_t \boldsymbol{M}_h\|^2 + \|(\boldsymbol{m}_h - \widehat{\boldsymbol{m}}_h)\lambda_h^n\|^2) \,\mathrm{d}t \\ + \tau_n^3 \|(\overline{\partial}\boldsymbol{m}_h^n - \overline{\partial}\widehat{\boldsymbol{m}}_h^n) \times \overline{\partial}^B \boldsymbol{m}_h^n\|^2 + \tau_n^3 \|(\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^n)\overline{\partial}\lambda_h^n\|^2 \Big).$$

We obtain the following theorem for variable time-steps (and non-changing mesh).

Theorem 29 (A posteriori error estimate)**.** *Let $\boldsymbol{m}$ be the exact solution of* (2.3.1) *and $\boldsymbol{m}_h^n$ the solution of* (2.3.2) *for $n = 2, \dots, N$ and the solution of* (2.3.7) *for $n = 1$. Assume $\tau = \tau_1 = \tau_2$ and $V_h^n = V_h$ for all $n = 0, \dots, N$. Then*

$$\|\boldsymbol{\nabla}(\boldsymbol{m} - \boldsymbol{m}_h)\|^2_{\mathcal{L}^\infty(t_0,t_N;\boldsymbol{\mathcal{L}}^2(\Omega))} \lesssim \mathcal{F}_1 + \mathcal{F}_2 + \widetilde{\mathcal{E}}_1 + \widetilde{\mathcal{E}}_2 + \widetilde{\mathcal{E}}_3 + \widetilde{\mathcal{E}}_4 + \widetilde{\mathcal{E}}_5 + \widetilde{\mathcal{E}}_6 + \widetilde{\mathcal{E}}_7 + \Lambda_1 + \widetilde{\Lambda}_2 + \widetilde{\Lambda}_3 \\ + \mathcal{P} + \mathcal{C}_1 + \widetilde{\mathcal{C}}_2 + \mathcal{Q}_1 + \widetilde{\mathcal{Q}}_2 + \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3,$$

where the (hidden) constant depends on α, $\|\boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}$, $\|\partial_t \boldsymbol{M}_h\|_{\boldsymbol{\mathcal{W}}^{1,\infty}}$, $\eta(\boldsymbol{m}_h^n; \mathcal{W}^{1,\infty})$ and $\|\lambda_h\|_{\mathcal{L}^\infty}$ and the estimators are given in Definition 18 and Definition 28.

CHAPTER 3

STABILITY

In this chapter, we consider an effective field that consists of the exchange field and an external field, i.e., $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \Delta \boldsymbol{m} + \boldsymbol{h}_{\mathrm{ext}}$. Multiplying the alternative form of the LLG equation (1.2.16) with $\partial_t \boldsymbol{m}$ and integrating over the time-space domain $\Omega \times (0,T)$ yields

$$\frac{1}{2}\ell_{\mathrm{ex}}^2 \|\boldsymbol{\nabla} \boldsymbol{m}(T)\|^2 + \alpha \int_0^T \|\partial_t \boldsymbol{m}\|^2 = \frac{1}{2}\ell_{\mathrm{ex}}^2 \|\boldsymbol{\nabla} \boldsymbol{m}(0)\|^2 + \int_0^T \int_\Omega \boldsymbol{h}_{\mathrm{ext}} \cdot \partial_t \boldsymbol{m}. \tag{3.0.1}$$

Thus, by Cauchy–Schwarz and Young's inequality, we obtain the inequality

$$\frac{1}{2}\ell_{\mathrm{ex}}^2 \|\boldsymbol{\nabla} \boldsymbol{m}(T)\|^2 + \frac{\alpha}{2} \int_0^T \|\partial_t \boldsymbol{m}\|^2 \leq \frac{1}{2}\ell_{\mathrm{ex}}^2 \|\boldsymbol{\nabla} \boldsymbol{m}(0)\|^2 + \frac{1}{2\alpha} \int_0^T \|\boldsymbol{h}_{\mathrm{ext}}\|^2. \tag{3.0.2}$$

Such inequalities are commonly referred to as energy inequalities, although they do not explicitly involve the total magnetic Gibbs free energy (1.2.3). However, their discrete counterparts provide a valuable robustness indicator for numerical schemes. For the tangent plane scheme analyzed in [7], a discrete energy inequality was established in [7, Prop. 3.3] for the uniform step-size BDF(k) method with $k = 1, 2$. The purpose of this section is to extend the analysis to the variable step-size BDF(k) (abbreviated as VS-BDF(k)) method with $k \in \{1, \dots, 5\}$. Moreover, the energy inequality can be employed as a tool to prove convergence without rates for a subsequence to a weak solution as $\tau, h \to 0$, see [13, 28, 58]. In close relation to the above inequality, integration by parts in (3.0.1) yields

$$\mathcal{E}_{\mathrm{tot}}(T) + \alpha \int_0^T \|\partial_t \boldsymbol{m}\|^2 = \mathcal{E}_{\mathrm{tot}}(0) - \int_0^T \int_\Omega \partial_t \boldsymbol{h}_{\mathrm{ext}} \cdot \boldsymbol{m},$$

where $\mathcal{E}_{\mathrm{tot}}(t) = \mathcal{E}_{\mathrm{ex}}(t) + \mathcal{E}_{\mathrm{ext}}(t) = \frac{1}{2}\ell_{\mathrm{ex}}^2 \|\boldsymbol{\nabla} \boldsymbol{m}(t)\|^2 - \int_\Omega \partial_t \boldsymbol{h}_{\mathrm{ext}}(t) \cdot \boldsymbol{m}(t)$ denotes the dimensionless total magnetic Gibbs free energy (1.2.3). If $\partial_t \boldsymbol{h}_{\mathrm{ext}} = 0$, this immediately implies the energy inequality $\mathcal{E}_{\mathrm{tot}}(T) \leq \mathcal{E}_{\mathrm{tot}}(0)$. This energy inequality is well studied in the literature and is usually included in the definition of a weak solution for the LLG equation [13, 14, 58, 86, 102], cf. Definition 6.

Typically, a discrete energy for parabolic equations using higher order BDF formulas for the discretization requires a restriction on the adjacent time-step ratios to hold. A classical result by Becker for the two-step BDF method requires the bound $\tau_n/\tau_{n-1} \leq (2 + \sqrt{13})/3 \approx 1.8685$ [32]. In [60], Emmrich improved the bound to 1.9104. For both of these results, the energy technique with a single multiplier was used. Recently, the bound for the BDF(2) method was slightly increased to 1.9398 by utilizing two multipliers in [9].

On the other hand, classical results on the zero-stability for variable step-size BDF schemes yield different bounds on the consecutive time-step ratios. Notably, Grigorieff obtained in

[77] the bound $1+\sqrt{2}$ for the zero-stability of the variable step-size BDF(2) method. In the subsequent work [39], Calvo et al. provide the zero-stability bound ≈ 1.476 for the variable step-size BDF(3) method and ≈ 1.101 for the variable step-size BDF(4) method. The bound for the variable step-size BDF(3) method was later improved to ≈ 1.501 in [78] through a spectral radius approach.

In Section 3.4 and 3.5, we establish G-stability of the variable step-size BDF(k) methods for $k = 2, 3$ with a single multiplier under a step-size ratio restriction that matches with the zero-stability bounds of [39, 77]. Subsequently, we obtain in Section 3.6 G-stability with a single multiplier for the variable step-size BDF(4) method with a step-size ratio bound of ≈ 1.150.

Finally, we obtain in Section 3.8 for the LLG equation a discrete energy inequality with the time-step ratios of Section 3.4-3.6 at the cost of a lower bound on the damping coefficient α. Similar to [7], we utilize the multiplier technique with a single multiplier combined with Dahlquist's G-stability to obtain a discrete energy inequality for the variable step-size BDF(k), $k = \{1, \dots 5\}$, method. Moreover, we concludes this chapter by extending the discrete energy inequality to an $\mathcal{L}^2$-bound for the discrete solution of the LLG equation utilizing the discrete orthogonal convolutions kernels introduced in Section 3.7. We begin with a description of the variable step-size BDF method and an overview of stability concepts for multistep methods.

3.1 Variable step-size BDF method

In this section, we aim to provide an introduction to the BDF methods with a focus on variable step-sizes. The BDF method is particularly valuable for stiff problems because its predictor, constructed through extrapolation, naturally aligns with the BDF formulation itself [55]. More precisely, the interpolation polynomial used to compute the solution implicitly at the current time-step t_n coincides with the predictor polynomial used for t_{n+1}. This characteristic is specific to BDF schemes and does not hold for other multistep methods for stiff differential equations. For an exemplary explanation, we refer to [55, Example 7.47]. In this work, we consider the BDF(k) method for $k \leq 5$ since only a single multiplier is required to ensure G-stability, which is often referred to as the Nevanlinna–Odeh multiplier (later denoted by η_k) in the literature due to [100]. However, for the six-step BDF method no Nevanlinna–Odeh multiplier exists and one requires at least a set of three multipliers [8]. Some results on the stability and an error analysis for the BDF(6) scheme can be found in [8, 46]. For some standard literature on time integration for ordinary differential equations, we refer to [55, 80, 81, 104], on which this section is partly based.

Consider the initial value problem on the interval $[0, T]$ with $T > 0$: Find $u \in C^1([0, T])$ such that

$$u'(t) = f(t, u(t)), \quad t \in (0, T), \tag{3.1.1}$$

with the initial condition $u(0) = u_0$. Further, let the time interval be partitioned as $0 =$

$t_0 < t_1 < \cdots < t_N = T$. Our goal is to approximate the solution at $t = t_n$ using previously computed values $u^{n-i} \approx u(t_{n-i})$ for $i = 1, \ldots, n$. A common approach is to differentiate the interpolation of the solution at t_n. Specifically, let $p \in \mathbb{P}^k$ be the interpolating polynomial satisfying

$$p(t_{n-i}) = u^{n-i}, \quad i = 0, \ldots, k. \tag{3.1.2}$$

Then, differentiating this polynomial yields the condition

$$p'(t_n) = f(t_n, u^n). \tag{3.1.3}$$

This is a representation of the BDF method, where we, in particular, do not require an equidistant partitioned time interval.

On an equidistant partition, it is also convenient to utilize backward differences to describe the polynomial, as well as the BDF scheme. Defining the discrete differences

$$\nabla_\tau^0 u^n := u^n, \qquad \nabla_\tau^\ell u^n := \nabla_\tau^{\ell-1} u^n - \nabla_\tau^{\ell-1} u^{n-1} \tag{3.1.4}$$

for $\ell \in \{1, \ldots, n\}$, we may express the polynomial (3.1.2) as

$$p(t) = \sum_{j=0}^{k} (-1)^j \binom{-s}{j} \nabla_\tau^j u^n \qquad \text{with} \qquad s = \frac{t - t_n}{\tau} \tag{3.1.5}$$

for $t \in \mathbb{R}$, where the value u^n has yet to be determined such that (3.1.3) holds [81, Cpt. 3 (1.19)]. By differentiating the polynomial, we obtain the linearly implicit formulas

$$\tau f(t_n, u^n) = \sum_{i=1}^{k} \frac{1}{i} \nabla_\tau^i u^n.$$

This can be rewritten to

$$\tau f(t_n, u^n) = \sum_{i=0}^{k} \delta_i u^{n-i}. \tag{3.1.6}$$

In the case of an equidistantly partitioned time interval, the coefficients δ_i correspond to the coefficients of the polynomial

$$\delta^k(\zeta) = \sum_{\ell=1}^{k} \frac{1}{\ell} (1 - \zeta)^\ell = \sum_{i=0}^{k} \delta_i \zeta^i. \tag{3.1.7}$$

Next, we want to present a technique to practically compute the coefficients δ_i for the BDF(k) method in the case of a non-uniform partition of the time interval. The formulas for the variable step-size BDF(k) scheme, $k = 1, \ldots, 4$, are given in [53, 121].

For notational convenience, we consider the partition

$$[0 = t_0, t_1, t_2, \dots, t_k] = [0, \tau_1, \tau_1 + \tau_2, \dots, \tau_1 + \cdots + \tau_k] \tag{3.1.8}$$

for the BDF(k) method for $k \geq 1$. Thus, we aim to approximate the solution at t_k using the previously computed values $u^i \approx u(t_i)$ for $i = 0, \dots, k-1$. The time interval may be shifted without loss of generality. Furthermore, we define the time-step ratio

$$\kappa_i \coloneqq \frac{\tau_i}{\tau_{i-1}} \tag{3.1.9}$$

for $i \geq 2$ for the following sections. Next, we define the abbreviation

$$\tau_{j,\dots,\ell} = \tau_j + \tau_{j+1} + \cdots + \tau_\ell \tag{3.1.10}$$

for $j < \ell$. We start by rewriting the BDF(k) method in the general form

$$\Phi^k(u;t) \coloneqq \delta_k u(t_0) + \delta_{k-1} u(t_1) + \cdots + \delta_1 u(t_{k-1}) + \delta_0 u(t) - (t - t_{k-1}) u'(t). \tag{3.1.11}$$

The numerical method is then defined by the condition $\Phi^k(u;t_k) = 0$, which determines the value of the solution at time t_k. We aim to compute the coefficients δ_i for the variable step-size BDF method. The coefficients can be obtained by enforcing the consistency condition

$$\Phi^k(u;t_k) = \mathcal{O}(\tau^{k+1}),$$

where $\tau = \max_{i=1,\dots,k} \tau_i$ denotes the largest step-size. However, consistency alone does not guarantee the convergence of multistep methods. In the case of a multistep method with a constant step-size, consistency combined with zero-stability is equivalent to convergence [55, Cpt. 7.1.3]. For variable step-sizes, zero-stability and an appropriate step-ratio bound is usually required in addition to consistency such that the method remains convergent [48, 77, 115]. Hence, by the Taylor expansion of $u(t_i)$ about t_k for $i = 0, \dots, k-1$, we obtain

$$u(t_i) = u(t_k) - \tau_{i,\dots,k} u'(t_k) + \frac{1}{2} \tau_{i,\dots,k}^2 u''(t_k) + \dots. \tag{3.1.12}$$

Substituting the Taylor expansion into (3.1.11) leads to the following system of equations for the coefficients

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & \tau_k & \tau_{k-1,k} & \dots & \tau_{1,\dots,k} \\ 0 & \tau_k^2 & \tau_{k-1,k}^2 & \dots & \tau_{1,\dots,k}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \tau_k^k & \tau_{k-1,k}^k & \dots & \tau_{1,\dots,k}^k \end{bmatrix} \begin{bmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \\ \vdots \\ \delta_k \end{bmatrix} = \begin{bmatrix} 0 \\ -\tau_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.1.13}$$

to ensure consistency. For example, if $k = 3$, we can compute the coefficients directly using

Cramer's rule by

$$\delta_1 = \frac{-\tau_{1,2,3}\tau_{2,3}}{\tau_{1,2}\tau_2}, \qquad \delta_2 = \frac{\tau_{1,2,3}\tau_3^2}{\tau_1\tau_2\tau_{2,3}}, \qquad \delta_3 = \frac{-\tau_{2,3}\tau_3^2}{\tau_1\tau_{1,2}\tau_{1,2,3}}. \tag{3.1.14}$$

The remaining value is then given by the first equation as $\delta_0 = -(\delta_1 + \delta_2 + \delta_3)$.

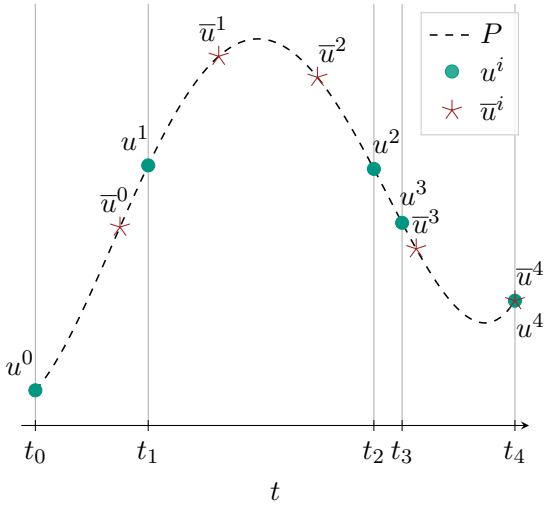


Figure 3.1.1: The dashed line shows the Lagrange interpolation polynomial P of $u_0, \ldots, u_4$, marked with green circles, as defined in (3.1.15). The values on the uniform time mesh $\overline{u}_1, \ldots, \overline{u}_4$, indicated with red stars, are obtained by evaluating the polynomial P according to (3.1.16). At t_4, the values of u_4 and $\overline{u}_4$ coincide.

However, the transposed Vandermonde matrix in (3.1.13) becomes ill-conditioned for large values of k. Thus, solving the linear equation system numerically (3.1.13) may be unreliable. Alternatively, we could project the recently computed values onto a locally uniform mesh and utilize the uniform step-size BDF(k) formulas. To do this, let us assume that the values $u^i = u(t_i)$ for $i = 0, \ldots, k-1$ are given and we want to compute the value u^k using the BDF(k) method for $k \geq 1$ with a time-step size τ_k. Further, let

$$L_j^n(t) \coloneqq \prod_{i=0, i\neq j}^{n} \frac{t - t_i}{t_j - t_i}$$

be the Lagrange basis polynomials and

$$P(t) \coloneqq P[u|t_0, \ldots, t_{k-1}](t) \coloneqq \sum_{j=0}^{k-1} u^j L_j^{k-1}(t) \tag{3.1.15}$$

be the interpolating polynomial of the known values u^i for $i = 0, \ldots, k-1$ in $\mathbb{P}_{k-1}$. Then, we begin by evaluating the values on the uniform time grid with a mesh size of τ_k by

$$\overline{u}^{k-i} \coloneqq P(t_k - i\,\tau_k), \qquad i = 1, \ldots, k. \tag{3.1.16}$$

Once these values are obtained, we can compute the value of u^k using the uniform time-step

size BDF(k) scheme with the values $\overline{u}^{k-i}$ for $i = 1, \dots, k$. In Figure 3.1.1, we illustrate an example of interpolation onto a uniform time mesh using the Lagrange interpolation polynomial (3.1.15).

3.2 Stability concepts for multistep methods

In this section, we briefly recall fundamental concepts for the stability analysis of uniform step-size multistep methods. This section is based on [80, 81].

Consider the ordinary differential equation $u'(t) = f(t, u(t))$ as introduced in (3.1.1). We begin by defining a general uniform step-size k-step multistep method as

$$\sum_{j=0}^{k} \alpha_j u^{n+j} = \tau \sum_{j=0}^{k} \beta_j f(t_{n+j}, u^{n+j}). \tag{3.2.1}$$

The associated *one-leg* method is given by

$$\sum_{j=0}^{k} \alpha_j u^{n+j} = \tau f\Big(\sum_{j=0}^{k} \beta_j t_{n+j}, \sum_{j=0}^{k} \beta_j u^{n+j} \Big). \tag{3.2.2}$$

In particular, for the BDF(k) method, we have $\beta_k = 1$ and $\beta_0 = \cdots = \beta_{k-1} = 0$. In this case, the multistep method coincides with its one-leg counterpart. The *generating polynomials* for the multistep method (3.2.1) are defined by

$$\rho(z) = \sum_{i=0}^{k} \alpha_i z^i, \qquad \sigma(z) = \sum_{i=0}^{k} \beta_i z^i. \tag{3.2.3}$$

To analyze the long-time behavior of a multistep method, we first investigate the zero-stability, which is closely linked to the generating polynomial ρ. This condition ensures that the numerical solution remains bounded as $n \to \infty$ when the multistep method is applied to the trivial test equation $u' = 0$.

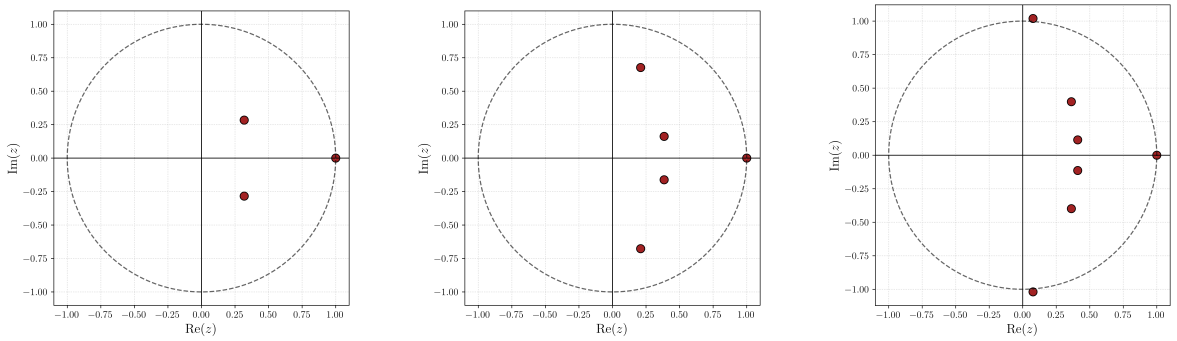
Definition 30 (Zero-stability)**.** *A multistep method* (3.2.1) *is called* zero-stable *if the roots of the generating polynomial ρ given by* (3.2.3) *satisfy the root conditions:*

(i) The roots of ρ lie on or within the unit circle.

(ii) The roots on the unit circle are simple.

Figure 3.2.1 illustrates the roots of the generating polynomial ρ for the BDF(k) method with $k = 3, 5, 7$. While each BDF scheme possesses a simple root on the unit circle, the illustration shows that for $k = 7$, there exist roots outside the unit circle, confirming that BDF(7) is unstable. In contrast, for $k \leq 6$, all roots satisfy the root conditions. This is summarized in the following theorem from [81, III.3; Thm. 3.4].

Theorem 31 (Zero-stability of BDF(k))**.** *The* BDF(k) *method is zero-stable for $k \leq 6$ and unstable for $k \geq 7$.*



(a) $k = 3$ (b) $k = 5$ (c) $k = 7$

Figure 3.2.1: Roots of the generating polynomial ρ for the uniform step-size BDF(k) method with $k = 3, 5, 7$.

While zero-stability is a fundamental requirement for the convergence of a uniform step-size multistep method, it only ensures stability as the step-size $\tau \to 0$. However, in particular for stiff differential equations, we require the method to remain stable for fixed step-sizes $\tau > 0$. This leads to the concept of A-stability, which studies the behavior of a method applied to the linear test problem $u' = \lambda u$.

Definition 32 (A-stability)**.** *A multistep method* (3.2.1) *is called* A-stable *if* $\mathbb{C}^- := \{z \in \mathbb{C} : \operatorname{Re} z \leq 0\} \subset S$*, where* S *is the stability region defined by*

$$S := \{z \in \mathbb{C} : z = \tau\lambda;\ \textit{the numerical solution of } u' = \lambda u \textit{ with step-size } \tau \textit{ is bounded}\}. \tag{3.2.4}$$

Instead of checking that the left half-plane is contained in the stability region, we can analyze the ratio of the generating polynomials ρ and σ given in (3.2.3). This algebraic approach is summarized in the following theorem from [80, V.1; Thm. 1.5].

Theorem 33 (A-stability condition)**.** *If the multistep method* (3.2.1) *is A-stable, then the generating polynomials* (3.2.3) *satisfy*

$$\operatorname{Re} \frac{\rho(z)}{\sigma(z)} > 0 \qquad \text{for } |z| > 1. \tag{3.2.5}$$

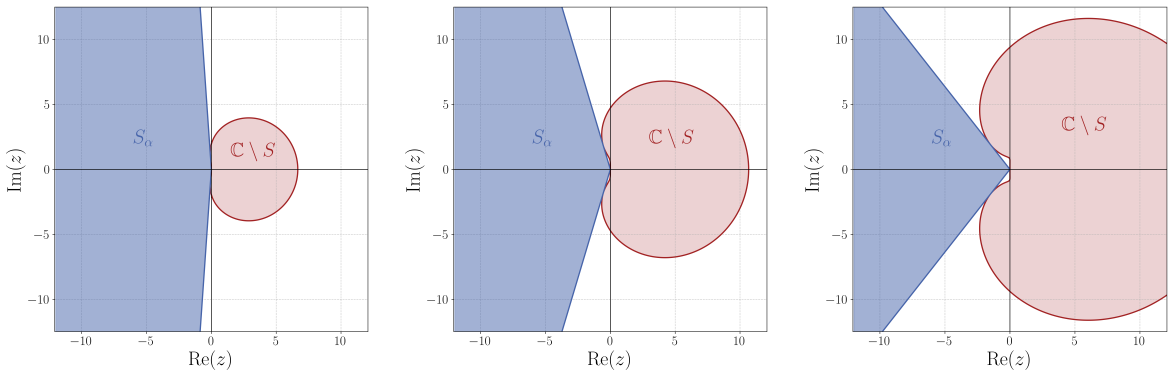
For irreducible multistep methods, the converse is also true.

Since higher-order multistep methods may fail to satisfy A-stability, we consider the following less restrictive stability condition.

Definition 34 ($A(\alpha)$-stability)**.** *A multistep method* (3.2.1) *with stability domain* S *is called* $A(\alpha)$*-stable for* $0 < \alpha < \pi/2$*, if*

$$S \supset S_\alpha := \{z \in \mathbb{C} : |\arg(-z)| \leq \alpha, z \neq 0\}. \tag{3.2.6}$$

In Figure 3.2.2, we depict the stability domain for the BDF(k) method for $k = 3, 4, 5$.



(a) $k = 3$ (b) $k = 4$ (c) $k = 5$

Figure 3.2.2: Stability domain S (3.2.4) and S_α (3.2.6) for the uniform step-size BDF(k) method with $k = 3, 4, 5$.

While these methods are not A-stable, they exhibit $A(\alpha)$-stability. The angle α decreases as k increases.

While A-stability is defined for the linear test problem $u' = \lambda u$, we consider now a right-hand side f that satisfies the one-sided Lipschitz condition

$$\operatorname{Re} \langle f(t, x) - f(t, y), x - y \rangle \leq \mu \|x - y\|^2. \tag{3.2.7}$$

To analyze the stability of multistep methods in this nonlinear setting, we introduce the history vector

$$U^n = [u^n, \ldots, u^{n-k+1}]^\top.$$

Then, we define with the inner product $\langle \cdot, \cdot \rangle$ the norm

$$\|U^n\|_G^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} g_{ij} \langle u^{n-k+i}, u^{n-k+j} \rangle,$$

where $G = (g_{ij})_{i,j=1,\ldots,k} \in \mathbb{R}^{k \times k}$ is a symmetric and positive definite matrix. With this framework, we introduce the concept of G-stability as follows, which considers the contractivity of the numerical solution.

Definition 35 (G-stability)**.** *A one-leg method* (3.2.2) *is called* G-stable*, if there exists a real, symmetric and positive definite matrix $G \in \mathbb{R}^{k \times k}$, such that for two numerical solutions $\{u^n\}$ and $\{\widetilde{u}^n\}$, we have*

$$\|U^n - \widetilde{U}^n\|_G \leq \|U^{n-1} - \widetilde{U}^{n-1}\|_G \tag{3.2.8}$$

for all step-sizes $\tau > 0$ and for all right-hand sides satisfying a one-sided Lipschitz condition (3.2.7) *with $\mu = 0$.*

In the context of Runge–Kutta methods, an analogous counterpart is given by the B-stability

[80, IV.12]. Since the test problem for A-stability $u' = \lambda u$ satisfies the one-sided Lipschitz condition (3.2.7) with $\mu = 0$, G-stability immediately implies A-stability. The converse was proven by Dahlquist in [49] yielding the equivalence of A- and G-stability, which is stated in the following theorem from [80, V.6; Thm. 6.7].

Theorem 36 (A-stability is equivalent to G-stability)**.** *If the generating polynomials ρ and σ given by* (3.2.3) *have no common divisor, then the multistep method* (3.2.1) *is A-stable if and only if the corresponding one-leg method* (3.2.2) *is G-stable.*

While the previous results establish a link between A- and G-stability, many practical schemes such as higher-order BDF methods only satisfy the weaker requirement of $A(\alpha)$-stability. Moreover, the assumption of a one-sided Lipschitz condition can be too restrictive for certain nonlinear problems. To address these limitations, we aim to extend the G-stability framework to a more general setting. To this end, we shift our focus to the generating function

$$\delta(z) = \frac{\rho(1/z)}{\sigma(1/z)}. \tag{3.2.9}$$

The function δ plays a central role in the multiplier technique introduced by Nevanlinna and Odeh in [100]. By utilizing this representation, we can investigate whether a condition similar to the A-stability condition (3.2.5) can be recovered for the broader class of $A(\alpha)$-stable methods. First, we introduce the concept of a multiplier.

Definition 37 (Multiplier)**.** *A rational function μ is called a* multiplier *for the generating function δ given by* (3.2.9)*, if $\mu \not\equiv \delta$ and*

$$\operatorname{Re} \frac{\delta(1/z)}{\mu(1/z)} > 0 \qquad \text{for } |z| > 1,$$

or, equivalently,

$$\operatorname{Re} \frac{\delta(z)}{\mu(z)} > 0 \qquad \text{for } |z| < 1. \tag{3.2.10}$$

The condition (3.2.10) generalizes the A-stability criterion (3.2.5) and acts as the starting point for deriving the generalized G-stability condition introduced in the next section.

The following example illustrates the generating functions for the BDF(2) method.

Example 38 (Generating functions for BDF(2))**.** *The uniform step-size* BDF(2) *method is given by*

$$\frac{3}{2}u^{n+2} - 2u^{n+1} + \frac{1}{2}u^n = \tau f(t_{n+2}, u^{n+2}).$$

Thus, we obtain the generating polynomials

$$\rho(z) = \frac{3}{2}z^2 - 2z + \frac{1}{2}, \qquad \sigma(z) = z^2.$$

By (3.2.9)*, we obtain*

$$\delta(z) = \frac{3}{2} - 2z + \frac{1}{2}z^2,$$

which coincides with (3.1.7) *introduced earlier.*

The following example provides an analysis of the eigenvalues for a linearized LLG equation. As the damping parameter α approaches zero, the eigenvalues shift toward the imaginary axis and may fall outside the stability regions of higher-order BDF methods, as illustrated in Figure 3.2.3.

Example 39 (Eigenvalues of a linearized LLG equation)**.** *We consider the alternative form of the LLG equation* (1.5.1) *linearized about* $\widehat{\boldsymbol{m}} = [0,0,1]^\top$ *with the effective field* $\boldsymbol{h}_{\text{eff}} = \ell_{\text{ex}}^2 \Delta \boldsymbol{m}$*, that is,*

$$\alpha \partial_t \boldsymbol{m} + \widehat{\boldsymbol{m}} \times \partial_t \boldsymbol{m} = \ell_{\text{ex}}^2 \boldsymbol{P}(\widehat{\boldsymbol{m}}) \Delta \boldsymbol{m}. \tag{3.2.11}$$

By explicit computations, we obtain

$$\alpha \begin{bmatrix} \partial_t m_x \\ \partial_t m_y \end{bmatrix} + \begin{bmatrix} -\partial_t m_y \\ \partial_t m_x \end{bmatrix} = \ell_{\text{ex}}^2 \begin{bmatrix} \Delta m_x \\ \Delta m_y \end{bmatrix} \tag{3.2.12}$$

and $\partial_t m_z = 0$*, where* $\boldsymbol{m} = [m_x, m_y, m_z]^\top$*. Introducing the complex-valued variable* $\psi^+ \coloneqq m_x + \mathrm{i}\, m_y$*, we obtain*

$$(\alpha + \mathrm{i}) \partial_t \psi^+ = \ell_{\text{ex}}^2 \Delta \psi^+. \tag{3.2.13}$$

Now, let λ_j*,* $j \in \mathbb{N}$*, denote an eigenvalue of the Laplacian* $\ell_{\text{ex}}^2 \Delta$ *(with appropriate boundary conditions), cf. [61, Sct. 6.5]. Then, we obtain the eigenvalues of* (3.2.13) *by*

$$\widetilde{\lambda}_j^+ = \frac{1}{\alpha + \mathrm{i}} \lambda_j = \frac{\alpha - \mathrm{i}}{1 + \alpha^2} \lambda_j \tag{3.2.14}$$

for $j \in \mathbb{N}$*. Analogously, defining* $\psi^- \coloneqq m_x - \mathrm{i}\, m_y$*, we obtain the eigenvalues*

$$\widetilde{\lambda}_j^- = \frac{1}{\alpha - \mathrm{i}} \lambda_j = \frac{\alpha + \mathrm{i}}{1 + \alpha^2} \lambda_j \tag{3.2.15}$$

for $j \in \mathbb{N}$*.*

Suppose the domain $(0, L)$ *is discretized uniformly with the nodes* $x_j = jh$ *for* $j = 1, \ldots, N-1$*, where* h *denotes the mesh size. Under homogeneous Dirichlet boundary conditions, the eigenvalues of the resulting discrete Laplacian, approximated by the second-order central difference quotient and scaled by* ℓ_{ex}^2*, are given by*

$$\lambda_{j,h} = -\frac{4\ell_{\text{ex}}^2}{h^2} \sin^2\left(\frac{\pi j}{2N}\right)$$

for $j = 1, 2, \ldots, N-1$, see [104, Sct. 8.9.4]. Following (3.2.14) *and* (3.2.15)*, the eigenvalues for the discretized linearized LLG equation* (3.2.11) *are then given by*

$$\widetilde{\lambda}_{j,h}^{\pm} = \frac{\alpha \mp \mathrm{i}}{1+\alpha^2}\lambda_{j,h}. \tag{3.2.16}$$

These eigenvalues are illustrated in Figure 3.2.3. As the damping coefficient $\alpha \to 0$, the eigenvalues approach the imaginary axis.

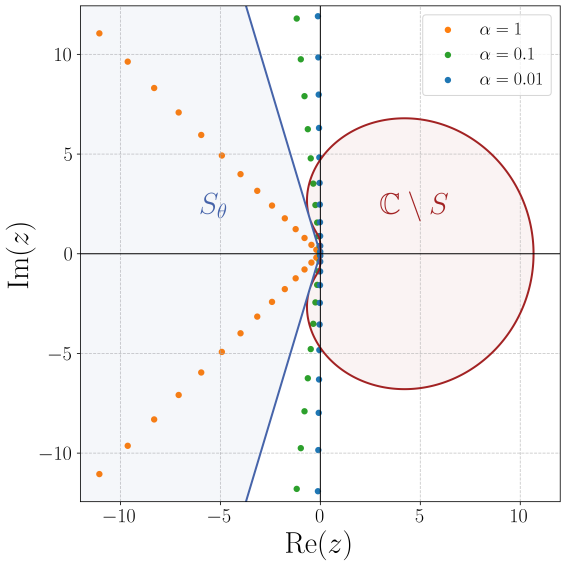


Figure 3.2.3: Illustration of the eigenvalues (3.2.16) for $\alpha \in \{1, 0.1, 0.01\}$ and the stability domain S (3.2.4) and S_θ (3.2.6) with the angle θ for the uniform step-size BDF(4) method, cf. Figure 3.2.2.

3.3 Multiplier technique

In this section, we want to recall the fundamental theorem on G-stability by Dahlquist [49], which is crucial in the later proof of the discrete energy inequality for the LLG equation. In addition, we review the multipliers associated with the uniform step-size BDF formulas. This section is mostly based on [7, Appendix A] and [100].

Lemma 40 (G-stability)**.** *Let δ and μ be polynomials of degree at most k (and at least one of them of degree k) that have no common divisor. Let V be a real vector space with the inner product $\langle \cdot, \cdot \rangle$. If*

$$\operatorname{Re} \frac{\delta(z)}{\mu(z)} > 0 \qquad \text{for } |z| < 1,$$

then there exists a positive definite matrix $G \in \mathbb{R}^{k \times k}$ with entries g_{ij}, $i, j \in \{1, \ldots, k\}$, such

that

$$\langle \sum_{i=0}^{k} \delta_i v^{k-i}, \sum_{j=0}^{k} \mu_j v^{k-j} \rangle \geq \sum_{i,j=1}^{k} g_{ij} \langle v^i, v^j \rangle - \sum_{i,j=1}^{k} g_{ij} \langle v^{i-1}, v^{j-1} \rangle \tag{3.3.1}$$

for $v^0, \dots v^k \in V$.

If we consider a discretization by a one-leg method (3.2.2) of $u'(t) = f(t, u(t))$, where f satisfies the one-sided Lipschitz condition (3.2.7) with $\mu = 0$, then the G-stability condition (3.3.1) immediately reduces to (3.2.8). The explicit construction of the matrix G is demonstrated in the following example for the BDF(2) method, cf. [80, V.6; Example 6.5].

Example 41 (G-matrix for BDF(2))**.** *For the second-order BDF method, we have*

$$\delta_0 = \frac{3}{2}, \qquad \delta_1 = -2, \qquad \delta_2 = \frac{1}{2}.$$

Thus the left-hand side of (3.3.1) *is given by*

$$lhs \coloneqq \langle \frac{3}{2}u^2 - 2u^1 + \frac{1}{2}u^0, u^2 \rangle = \frac{3}{2}\|u^2\| - 2\langle u^1, u^2 \rangle + \frac{1}{2}\langle u^0, u^2 \rangle$$

The right-hand side of (3.3.1) *is given by*

$$rhs \coloneqq \Big(g_{11}\|u^1\|^2 + 2g_{12}\langle u^1, u^2 \rangle + g_{22}\|u^2\|^2 \Big) - \Big(g_{11}\|u^0\|^2 + 2g_{12}\langle u^0, u^1 \rangle + g_{22}\|u^1\|^2 \Big).$$

Defining

$$S \coloneqq \frac{1}{4}\|u^2 - 2u^1 + u^0\|^2 = \frac{1}{4}\Big(\|u^2\|^2 + 4\|u^1\|^2 + \|u^0\|^2 - 4\langle u^1, u^2 \rangle + 2\langle u^0, u^2 \rangle - 4\langle u^0, u^1 \rangle \Big),$$

we obtain G *by comparing the coefficients of* $lhs = rhs + S$*. This yields*

$$G = \frac{1}{4}\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix}.$$

The coefficients of the polynomial μ are commonly referred to as multipliers. As an example, we can consider the uniform BDF formulas, whose coefficients correspond to the coefficients of the polynomial

$$\delta^k(z) = \sum_{\ell=1}^{k} \frac{1}{\ell}(1-z)^\ell = \sum_{i=0}^{k} \delta_i z^i. \tag{3.3.2}$$

A possible multiplier is given by $\mu^k(z) = 1 - \eta_k z$ for $k \leq 5$, summarized in the following theorem from [100]. Note that due to the A-stability of the BDF(2) scheme, no multiplier is needed to ensure the G-stability.

Lemma 42 (Multiplier for uniform step-size BDF)**.** *Let $k \leq 5$ and δ^k be given in* (3.3.2)*. Then, there exists $0 \leq \eta_k < 1$ such that*

$$\operatorname{Re} \frac{\delta^k(z)}{1 - \eta_k z} > 0 \qquad \text{for } |z| < 1.$$

The smallest possible values of η_k are

$$\eta_1 = \eta_2 = 0, \quad \eta_3 = 0.0836, \quad \eta_4 = 0.2878, \quad \eta_5 = 0.8160.$$

3.4 A multiplier for the variable step-size BDF(2) scheme

In this section, we consider the variable step-size BDF(2) scheme on the partition

$$[0, t_1, t_2] = [0, \tau_1, \tau_1 + \tau_2],$$

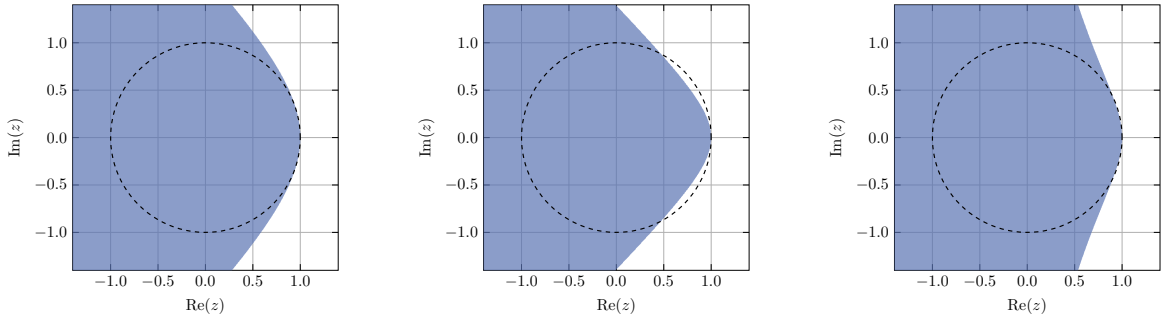
cf. (3.1.8). Since we require only a single time-step ratio to describe the variable step-size BDF(2) method, we define $\kappa := \tau_2/\tau_1$. Then, the characteristic polynomial of the variable step-size BDF(2) method is described by

$$\delta(z) := \sum_{i=0}^{2} \delta_i z^i, \tag{3.4.1}$$

where the coefficients depending on the step-size ratio κ are given by

$$\delta_0(\kappa) := \frac{1 + 2\kappa}{1 + \kappa}, \quad \delta_1(\kappa) := -(1 + \kappa), \quad \delta_2(\kappa) := \frac{\kappa^2}{1 + \kappa}. \tag{3.4.2}$$

This section is mostly based on [36].



(a) $\kappa = 1$, $\eta = 0$ (b) $\kappa = 1.4$, $\eta = 0$ (c) $\kappa = 1.4$, $\eta = 0.3$

Figure 3.4.1: Stability region $\operatorname{Re}(\delta(z)/(1 - \eta z)) > 0$ for variable step-size BDF(2) method for varying parameters κ and η. The coefficients of the polynomial δ are given in (3.4.2). If $\kappa > 1$, the complex sphere $|z| < 1$ is not completely contained. The multiplier $\eta > 0$ increases the stability region.

The uniform step-size BDF(2) scheme is G-stable without requiring any multiplier, as stated in Lemma 42 and illustrated in Figure 3.4.1. However, as soon as the time-step ratio $\kappa > 1$,

we need to introduce a multiplier $\eta > 0$ to preserve the G-stability. Further representative examples of the stability region $\mathrm{Re}(\delta(z)/(1-\eta z)) > 0$ with different values for κ and η are visualized in Figure 3.4.1.

The bound of $1+\sqrt{2}$ for the step-size ratio, which we obtain in the following lemma, dates back to [77] for the zero-stability of the variable step-size BDF(2) method but was proved here using the multiplier technique from [100].

Lemma 43 (Multiplier for variable step-size BDF(2))**.** *Let $\delta(z) = \sum_{i=0}^{2} \delta_i z^i$ be the characteristic* BDF(2) *polynomial* (3.4.1)*, where the coefficients δ_i for $i = 0, 1, 2$ are given in* (3.4.2)*. Furthermore, let $\kappa \in (0, 1+\sqrt{2})$. Then, there exists a multiplier $\eta \in [0, 1)$, such that*

$$\mathrm{Re}\,\frac{\delta(z)}{1-\eta z} > 0 \qquad \text{for } |z| < 1.$$

Moreover, the multiplier satisfies

$$\eta > \frac{3\kappa^2 - 2\kappa - 1}{\kappa^2 + 2\kappa + 1}.$$

Proof. We see immediately that δ has two roots $z_1 = 1$ and $z_2(\kappa) = (1+2\kappa)/\kappa^2 > 1$ as long as $\kappa < 1+\sqrt{2}$. Thus, we have

$$\frac{\delta(z)}{1-\eta z} = \frac{\delta_2(z-1)(z-z_2)}{1-\eta z} = \frac{\delta_2(z-1)(z-z_2)(1-\eta\overline{z})}{|1-\eta z|^2}.$$

Clearly, the choice $\eta = 1/z_2$ would suffice to ensure $\mathrm{Re}(\delta(z)/(1-\eta z)) > 0$ for $|z| < 1$. However, we obtain a sharper bound by resolving the inequality

$$\mathrm{Re}\,((z-1)(z-z_2)(1-\eta\overline{z})) > 0 \qquad \text{for } |z| < 1.$$

Rewriting $z = x + \mathrm{i}y$, $x, y \in \mathbb{R}$, and considering the edge case $|z| = 1$ leads to the quadratic inequality (the cubic terms cancel)

$$2x^2 - (1+\eta)(1+z_2)x + (1+\eta)z_2 + \eta - 1 > 0$$

for $x \in (-1, 1)$. Since $(1-x)(3-z_2) < -2x^2 + x(1+z_2) + (1+z_2)$ for $x \in (-1, 1)$, this can be fulfilled for $0 < \kappa < 1+\sqrt{2}$ if

$$\eta > \frac{3-z_2}{z_2+1} = \frac{3\kappa^2 - 2\kappa - 1}{\kappa^2 + 2\kappa + 1} =: g(\kappa).$$

Since $g(\kappa) \le 0$ for $\kappa \in (0, 1]$, we may choose $\eta = 0$. Since $0 < g(\kappa) < 1$ for $\kappa \in (1, 1+\sqrt{2})$, we always find $\eta < 1$ in this case. □

3.5 A multiplier for the variable step-size BDF(3) scheme

In this section, we consider the variable step-size BDF(3) scheme on the partition

$$[0, t_1, t_2, t_3] = [0, \tau_1, \tau_1 + \tau_2, \tau_1 + \tau_2 + \tau_3],$$

cf. (3.1.8), and use the ratio of consecutive time-steps $\kappa_\ell = \tau_\ell / \tau_{\ell-1} > 0$ for $\ell \geq 2$ as defined in (3.1.9). Since the uniform step-size BDF(3) scheme already requires a multiplier $\eta = 0.0836$ for G-stability, see Lemma 42, any choice of $\kappa_\ell > 1$ likewise necessitates a multiplier $\eta > 0$ to preserve G-stability. The variable step-size BDF(3) scheme is described by the polynomial

$$\delta(z) = \sum_{i=0}^{3} \delta_i z^i, \tag{3.5.1}$$

where the coefficients are given by

$$\begin{aligned}
\delta_0(\kappa_3, \kappa_2) &= \frac{\kappa_3\kappa_2(\kappa_3+1) + (2\kappa_3+1)(\kappa_3\kappa_2+\kappa_2+1)}{(\kappa_3+1)(\kappa_3\kappa_2+\kappa_2+1)}, \\
\delta_1(\kappa_3, \kappa_2) &= \frac{-\kappa_3^2\kappa_2 - 2\kappa_3\kappa_2 - \kappa_3 - \kappa_2 - 1}{\kappa_2+1}, \\
\delta_2(\kappa_3, \kappa_2) &= \frac{\kappa_3^2(\kappa_3\kappa_2+\kappa_2+1)}{\kappa_3+1}, \\
\delta_3(\kappa_3, \kappa_2) &= \kappa_3^2\kappa_2^3 \frac{-\kappa_3 - 1}{(\kappa_2+1)(\kappa_3\kappa_2+\kappa_2+1)}
\end{aligned} \tag{3.5.2}$$

for $\kappa_3, \kappa_2 > 0$. These formulas can also be found in [53, App. A] or [92, (1.5)]. Next, to conveniently express the roots of the characteristic BDF(3) polynomial $\delta(z) = \sum_{i=0}^{3} \delta_i z^i$, we define the polynomials

$$\begin{aligned}
g(x,y) &= x^3y^2 + 4x^2y^2 + 2x^2y + 3xy^2 + 3xy + x, \\
h(x,y) &= x^6y^4 + 8x^5y^4 + 4x^5y^3 - 12x^4y^5 + 10x^4y^4 + 22x^4y^3 + 6x^4y^6 \\
&\quad - 40x^3y^5 - 24x^3y^4 + 28x^3y^3 + 20x^3y^6 + 4x^3y - 48x^6y^5 \\
&\quad - 59x^6y^4 - 2x^6y^3 + 15x^6y^6 + 6x^6y + x^6 - 24xy^5 - 40xy^4 \\
&\quad - 16xy^3 - 4y^5 - 8y^4 - 4y^3.
\end{aligned} \tag{3.5.3}$$

It is easy to check that $z_0 = 1$ is a root of δ. The other roots of δ are given by

$$\begin{aligned}
z_1(\kappa_3, \kappa_2) &= \frac{g(\kappa_3, \kappa_2) + \sqrt{h(\kappa_3, \kappa_2)}}{2\kappa_3\kappa_2^3(\kappa_3^2 + 2\kappa_3 + 1)}, \\
z_2(\kappa_3, \kappa_2) &= \frac{g(\kappa_3, \kappa_2) - \sqrt{h(\kappa_3, \kappa_2)}}{2\kappa_3\kappa_2^3(\kappa_3^2 + 2\kappa_3 + 1)}.
\end{aligned} \tag{3.5.4}$$

A visual representation of the roots for $\kappa_2 = 1.6$ and $\kappa_3 = 1.3$ is given in Figure 3.5.1. The following lemma establishes that the roots can be used to derive a lower bound for the Nevanlinna–Odeh multiplier for the variable step-size BDF(3) method.

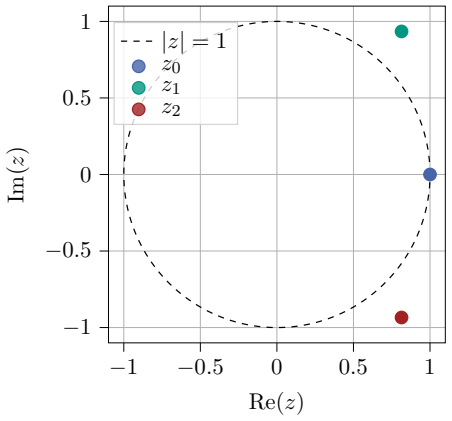


Figure 3.5.1: The roots (3.5.4) of the characteristic variable step-size BDF(3) polynomial (3.5.1) for $\kappa_2 = 1.6$ and $\kappa_3 = 1.3$.

Lemma 44 (Multiplier for variable step-size BDF(3))**.** *Let* $\delta(z) = \sum_{i=0}^{3} \delta_i z^i$ *be the characteristic* BDF(3) *polynomial* (3.5.1)*, where the coefficients* δ_i *for* $i = 0, \dots, 3$*, which depend on the step-size ratios* κ_3 *and* κ_2*, are given in* (3.5.2)*. Assume that the roots* z_1 *and* z_2 *given in* (3.5.4) *of* δ *lie outside the complex sphere, i.e.,* $|z_1| \geq 1$ *and* $|z_2| \geq 1$*. Then, there exists a multiplier* $\eta > 0$ *such that*

$$\operatorname{Re} \frac{\delta(z)}{1 - \eta z} > 0 \qquad \text{for } |z| < 1. \tag{3.5.5}$$

Moreover, let $z = x + \mathrm{i}y$ *with* $x, y \in \mathbb{R}$ *and define*

$$\mu(z) \coloneqq \mu(z, \kappa_3, \kappa_2) \coloneqq (z - 1)\big(z - z_1(\kappa_3, \kappa_2)\big)\big(z - z_2(\kappa_3, \kappa_2)\big). \tag{3.5.6}$$

If $\operatorname{Re}(\mu(z))x + \operatorname{Im}(\mu(z))y > 0$ *holds for all* z *with* $|z| = 1$*, then the multiplier satisfies*

$$\eta > \max_{z \in \mathbb{C} \,:\, |z| = 1} \frac{\operatorname{Re}(\mu(z))}{x \operatorname{Re}(\mu(z)) + y \operatorname{Im}(\mu(z))}. \tag{3.5.7}$$

Proof. By factoring the polynomial δ in terms of its roots and multiplying by $(1-\eta\overline{z})/(1-\eta\overline{z}) = 1$, we obtain

$$\frac{\delta(z)}{1 - \eta z} = \frac{\delta_3 (z-1)(z-z_1)(z-z_2)(1-\eta\overline{z})}{|1 - \eta z|^2} = \frac{\delta_3 \mu(z)(1-\eta\overline{z})}{|1 - \eta z|^2}.$$

Hence, we need to resolve the inequality

$$\operatorname{Re}\{\delta_3 \mu(z)(1 - \eta\overline{z})\} > 0$$

for $|z| < 1$ or, equivalently, since $\delta_3(\kappa_3, \kappa_2)$ given in (3.5.2) satisfies $\delta_3 < 0$ for $\kappa_3, \kappa_2 > 0$,

$$\operatorname{Re}\{\mu(z)(1 - \eta\overline{z})\} < 0$$

for $|z| < 1$. Thus, with $z = x + \mathrm{i}y$, $x, y \in \mathbb{R}$, we obtain

$$\begin{aligned}\operatorname{Re}\{\mu(z)(1-\eta\overline{z})\} &= \operatorname{Re}(\mu(z))(1-\eta x) - \operatorname{Im}(\mu(z))\eta y \\ &= -\eta(x\operatorname{Re}(\mu(z)) + y\operatorname{Im}(\mu(z))) + \operatorname{Re}(\mu(z)).\end{aligned}$$

Considering the edge case $|z| = 1$, we conclude with $\operatorname{Re}(\mu(z))x + \operatorname{Im}(\mu(z))y > 0$

$$\eta > \max_{z\in\mathbb{C}\,:\,|z|=1} \frac{\operatorname{Re}(\mu(z))}{x\operatorname{Re}(\mu(z)) + y\operatorname{Im}(\mu(z))}$$

since all roots of δ lie outside the complex sphere. □

Although Lemma 44 provides a Nevanlinna–Odeh multiplier $\eta > 0$, we are only interested in multipliers $\eta < 1$. Utilizing (3.5.7), we compute the combinations of κ_3 and κ_2 such that we obtain a multiplier $\eta < 1$. Moreover, we have to ensure that the roots (3.5.4) of the characteristic BDF(3) polynomial (3.5.1) lie outside the complex sphere. The results are given in Figure 3.5.2, where we computed the results numerically. The region of κ_2-κ_3-combinations guaranteeing $\eta < 1$ is indicated in blue. Further, the red region shows where $|z_2| < 1$.

Next, to compute an upper bound $\kappa^\star$, such that the variable step-size BDF(3) scheme is G-stable for $\kappa_3, \kappa_2 \in (0, \kappa^\star)$ with a multiplier $\eta < 1$, we define

$$f_0(\kappa) \coloneqq \max_{z\in\mathbb{C}\,:\,|z|=1} \left(\frac{\operatorname{Re}(\mu(z,\kappa,\kappa))}{\operatorname{Re}(z)\operatorname{Re}(\mu(z,\kappa,\kappa)) + \operatorname{Im}(z)\operatorname{Im}(\mu(z,\kappa,\kappa))} \right) - 1$$

for $\kappa > 0$, where μ is defined in (3.5.6). The key idea is that the smallest positive root of f_0 corresponds to $\kappa^\star$. To determine this value, we apply a root-finding algorithm, such as bisection or Brent's method [37, Cpt. 3-4], to obtain the smallest positive root of $\kappa^\star \approx 1.476$. Consequently, we find a multiplier $\eta < 1$ such that the variable step-size BDF(3) scheme is G-stable for $\kappa_3, \kappa_2 \in (0, \kappa^\star)$, as indicated by the dashed lines in Figure 3.5.2. This bound is well known for the BDF(3) scheme. For instance, Calvo et al. prove that the variable step-size BDF(3) scheme is zero-stable for the same value $\kappa^\star \approx 1.476$ in [39]. The same bound reappears later in Chapter 3.7 on discrete orthogonal convolution kernels.

Remark 45. *We may also increase the bound on κ_3 and κ_2 to ensure G-stability for the variable step-size* BDF(3) *with a multiplier $\eta < 1$ by freezing the time-step after coarsening. Thus, after switching from a time-step τ to $\kappa\tau$, two steps are taken with this new time-step size $\kappa\tau$. This leads to two scenarios that need to be analyzed: (I) $\tau \to \tau \to \kappa\tau$ and (II) $\tau \to \kappa\tau \to \kappa\tau$. Thus, there is only one parameter involved. In the first case (I), the* BDF(3) *coefficients* (3.5.2) *simplify to*

$$\begin{aligned}\delta_0(\kappa) &= 1 + \frac{\kappa}{1+\kappa} + \frac{\kappa}{2+\kappa}, & \delta_1(\kappa) &= -\frac{1}{2}(1+\kappa)(2+\kappa), \\ \delta_2(\kappa) &= \kappa^2\frac{2+\kappa}{1+\kappa}, & \delta_3(\kappa) &= -\frac{1}{2}\kappa^2\frac{1+\kappa}{2+\kappa},\end{aligned}$$

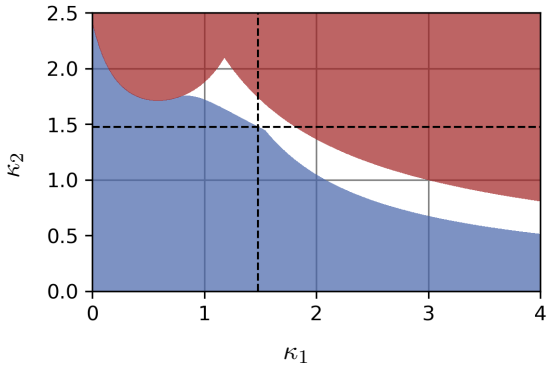


Figure 3.5.2: The blue region indicates where the variable step-size BDF(3) scheme is G-stable, i.e., there exists a multiplier $\eta < 1$ such that (3.5.5) is satisfied. The red region indicates where $|z_2| < 1$, i.e., there exists no multiplier for the variable step-size BDF(3) scheme. We note that $|z_2| \leq |z_1|$ and, in the case $\operatorname{Im}(z_2) \neq 0$, $|z_2| = |z_1|$, cf. (3.5.4). The dashed lines indicate $\kappa_2 = \kappa^\star \approx 1.476$ and $\kappa_3 = \kappa^\star$. Thus, the variable step-size BDF(3) scheme is G-stable for $\kappa_3, \kappa_2 \in (0, \kappa^\star)$ with a multiplier $\eta < 1$.

and, in the second case (II), to

$$\delta_0(\kappa) = \frac{3}{2} + \frac{\kappa}{1+2\kappa}, \qquad \delta_1(\kappa) = -2\frac{1+2\kappa}{1+\kappa},$$
$$\delta_2(\kappa) = \frac{1}{2}(1+2\kappa), \qquad \delta_3(\kappa) = -\frac{2\kappa^3}{(1+\kappa)(1+2\kappa)}.$$

Using the lower bound (3.5.7) *for the multiplier η given in Lemma 44, we compute exemplary values for case (I). The results are as follows:*

- *For $\kappa = 1.4$, we find $\eta \approx 0.572$.*
- *For $\kappa = 1.6$, we find $\eta \approx 0.833$.*
- *To determine the largest $\kappa_I^\star$ such that $\eta < 1$, we define*

$$f_I(\kappa) := \max_{z \in \mathbb{C} : |z|=1} \left(\frac{\operatorname{Re}(\mu(z,\kappa,1))}{\operatorname{Re}(z)\operatorname{Re}(\mu(z,\kappa,1)) + \operatorname{Im}(z)\operatorname{Im}(\mu(z,\kappa,1))} \right) - 1$$

and apply again the root finding algorithm of Brent [37, Cpt. 3-4] to compute the smallest positive root of f_I, yielding $\kappa_I^\star \approx 1.718$.

For the second case (II), we obtain the exemplary values:

- *For $\kappa = 1.4$, we find $\eta \approx 0.106$.*
- *For $\kappa = 2.0$, we find $\eta \approx 0.316$.*

- *To determine the largest $\kappa_{II}^\star$ such that $\eta < 1$, we define*

$$f_{II}(\kappa) := \max_{z\in\mathbb{C}\,:\,|z|=1} \left(\frac{\operatorname{Re}(\mu(z,1,\kappa))}{\operatorname{Re}(z)\operatorname{Re}(\mu(z,1,\kappa)) + \operatorname{Im}(z)\operatorname{Im}(\mu(z,1,\kappa))} \right) - 1$$

and apply again the root-finding algorithm of Brent [37, Cpt. 3-4] to compute the smallest positive root of f_{II}, yielding $\kappa_{II}^\star \approx 2.059$.

Naturally, the question arises what maximal coarsening we can achieve across two time-steps while ensuring G-stability with a multiplier $\eta < 1$. Let us consider the time-step sequence τ, τ, τ. Coarsening once yields the sequence $\tau \to \tau \to \kappa_2\tau$. Coarsening again, the sequence gets to $\tau \to \kappa_2\tau \to \kappa_3\kappa_2\tau$, leading to a total coarsening of $\kappa_3\kappa_2$. For instance, coarsening twice with $\kappa^\star \approx 1.476$ yields a total coarsening of $(\kappa^\star)^2 \approx 2.178$. In Figure 3.5.3 we illustrate that we find κ_2-κ_3-combinations, such that we achieve a higher coarsening than $(\kappa^\star)^2$. Further, we find the optimal two-step coarsening values of $\kappa_2 \approx 1.451$ and $\kappa_3 \approx 1.527$, yielding a total of $\kappa_2 \cdot \kappa_3 \approx 2.216$.

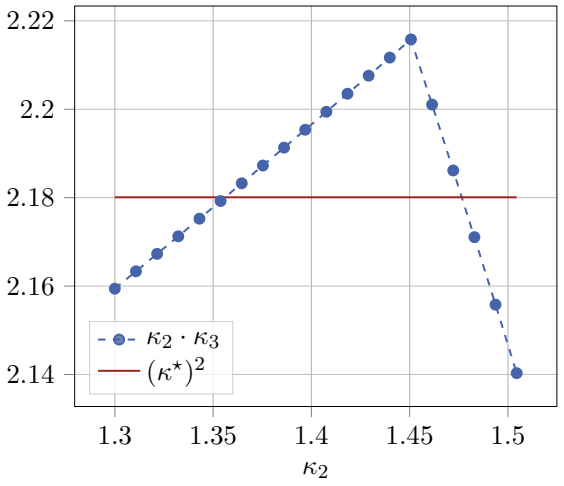


Figure 3.5.3: The red line indicates $(\kappa^\star)^2$ with $\kappa^\star \approx 1.476$. On the horizontal axis, we denote the value of the first coarsening step κ_2. The second coarsening step is chosen, such that the variable step-size BDF(3) method remains G-stable with a multiplier $\eta < 1$. The blue dots represent the multiplication of the first and second coarsening step $\kappa_2 \cdot \kappa_3$. The optimal combination is given by $\kappa_2 \approx 1.451$ and $\kappa_3 \approx 1.527$.

3.6 A multiplier for the variable step-size BDF(4) scheme

In this section, we consider the variable step-size BDF(4) scheme on the partition

$$[0, t_1, t_2, t_3, t_4] = [0, \tau_1, \tau_1 + \tau_2, \tau_1 + \tau_2 + \tau_3, \tau_1 + \tau_2 + \tau_3 + \tau_4],$$

cf. (3.1.8), and use the ratio of consecutive time-steps $\kappa_\ell = \tau_\ell/\tau_{\ell-1} > 0$ for $\ell \geq 2$ as defined in (3.1.9). To conveniently express the coefficients of the variable step-size BDF(4) polynomial,

we define the polynomials

$$\begin{aligned} A(x,y) &\coloneqq xy + y + 1, \\ B(x,y,z) &\coloneqq xyz + yz + z + 1 \end{aligned}$$

for $x, y, z \in \mathbb{R}$. Utilizing these polynomials, we can further define

$$\begin{aligned} d_0(x,y,z) &\coloneqq 1 + \frac{x}{x+1} + \frac{xy}{A(x,y)} + \frac{xyz}{B(x,y,z)}, \\ d_1(x,y,z) &\coloneqq -\frac{(x+1)A(x,y)\,B(x,y,z)}{(y+1)\,A(y,z)}, \\ d_2(x,y,z) &\coloneqq \frac{x^2 A(x,y)\,B(x,y,z)}{(x+1)(z+1)}, \\ d_3(x,y,z) &\coloneqq -\frac{x^2y^3(x+1)\,B(x,y,z)}{(y+1)\,A(x,y)}, \\ d_4(x,y,z) &\coloneqq \frac{x^2y^3z^4(x+1)A(x,y)}{(z+1)\,A(y,z)\,B(x,y,z)} \end{aligned} \tag{3.6.1}$$

for $x, y, z \in \mathbb{R}$. Obviously, we have $d_0, d_2, d_4 > 0$ and $d_1, d_3 < 0$ for $x, y, z > 0$. Then, the characteristic polynomial of the variable step-size BDF(4) scheme is given by

$$\delta(z) = \sum_{i=0}^{4} \delta_i z^i, \tag{3.6.2}$$

where the coefficients are obtained from the polynomials given in (3.6.1) as

$$\delta_i(\kappa_4, \kappa_3, \kappa_2) \coloneqq d_i(\kappa_4, \kappa_3, \kappa_2) \tag{3.6.3}$$

for $i = 0, \dots, 4$ and $\kappa_4, \kappa_3, \kappa_2 > 0$. These correspond to the coefficients given in [53, App. A].

To express the roots of the characteristic polynomial of the BDF(4) polynomial (3.6.2), we again define suitable polynomials

$$\begin{aligned} a(x,y,z) &\coloneqq (z+1)\widetilde{a}(x,y,z), \\ b(x,y,z) &\coloneqq y^3z^4(x+1)^2(y+1)A(x,y)^2, \\ c(x,y,z) &\coloneqq \widetilde{c}(x,y,z), \\ d(x,y,z) &\coloneqq z^4(y+1)A(x,y)^2, \\ e(x,y,z) &\coloneqq (z+1)A(y,z)\widetilde{e}(x,y,z), \\ f(x,y,z) &\coloneqq x^2y^3z^4(x+1)^2A(x,y)^2 \end{aligned}$$

for $x, y, z \in \mathbb{R}$. For the polynomials $\widetilde{a}$, $\widetilde{c}$ and $\widetilde{e}$ we did not find a suitable factorization. However, the precise definitions are given in Appendix A.1. It is evident that $a, b, c, d, e, f > 0$

for $x, y, z > 0$. Furthermore, we define the function $g : \mathbb{R}^6 \to \mathbb{C}$ by

$$g(a,b,c,d,e,f) \coloneqq \left(\sqrt{ -\left(\frac{c^2}{d^2} - \frac{3a}{b}\right)^3 + \left(-\frac{27e}{2f} + \frac{9ca}{2db} - \frac{c^3}{d^3}\right)^2 } - \frac{27e}{2f} + \frac{9ca}{2db} - \frac{c^3}{d^3} \right)^{1/3}.$$

For convenience, we define the short-hand notation

$$a_\kappa \coloneqq a(\kappa_4, \kappa_3, \kappa_2), \ \ldots, \ f_\kappa \coloneqq f(\kappa_4, \kappa_3, \kappa_2),$$

where the subscript κ denotes the dependence on $\kappa_4, \kappa_3, \kappa_2$. Further, for readability, we define

$$g_\kappa \coloneqq g(a_\kappa, \ldots, f_\kappa).$$

Then the roots z_i, $i = 0, 1, 2, 3$, of the BDF(4) polynomial are given by $z_0 = 1$,

$$\begin{aligned}
z_1(\kappa_4,\kappa_3,\kappa_2) &= \frac{\frac{3a_\kappa}{b_\kappa} - \frac{c_\kappa^2}{d_\kappa^2}}{3g_\kappa} - \frac{g_\kappa}{3} + \frac{c_\kappa}{3d_\kappa}, \\
z_2(\kappa_4,\kappa_3,\kappa_2) &= \frac{\frac{3a_\kappa}{b_\kappa} - \frac{c_\kappa^2}{d_\kappa^2}}{3\left(-\frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}\right) g_\kappa} - \frac{1}{6}\left(-1 - \mathrm{i}\sqrt{3}\right) g_\kappa + \frac{c_\kappa}{3d_\kappa}, \\
z_3(\kappa_4,\kappa_3,\kappa_2) &= \frac{\frac{3a_\kappa}{b_\kappa} - \frac{c_\kappa^2}{d_\kappa^2}}{3\left(-\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}\right) g_\kappa} - \frac{1}{6}\left(-1 + \mathrm{i}\sqrt{3}\right) g_\kappa + \frac{c_\kappa}{3d_\kappa}
\end{aligned} \tag{3.6.4}$$

for $\kappa_4, \kappa_3, \kappa_2 > 0$. A visual representation of the roots for $\kappa_2 = \kappa_3 = \kappa_4 = 1.15$ is given in Figure 3.6.1.

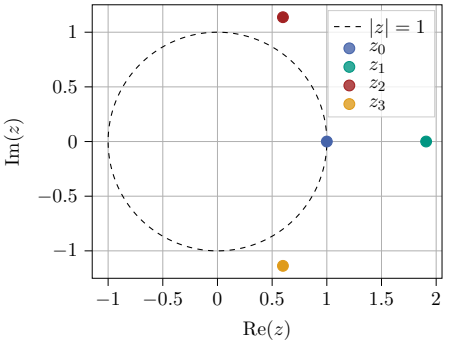


Figure 3.6.1: The roots (3.6.4) of the characteristic variable step-size BDF(4) polynomial (3.6.2) for $\kappa_2 = \kappa_3 = \kappa_4 = 1.15$.

Noticing that the roots of $\omega^3 = 1$ are given by

$$\omega_1 = 1, \qquad \omega_2 = -\frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}, \qquad \omega_3 = -\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2},$$

we may rewrite the roots as

$$z_i(\kappa_4,\kappa_3,\kappa_2) = \frac{\frac{3a_\kappa}{b_\kappa} - \frac{c_\kappa^2}{d_\kappa^2}}{3\omega_i g_\kappa} - \frac{1}{3}\omega_i g_\kappa + \frac{c_\kappa}{3d_\kappa}$$

for $i = 1, 2, 3$. Similar to Lemma 44 for the variable step-size BDF(3) method, we establish in the following lemma that the roots can be used to derive a lower bound for the Nevanlinna–Odeh multiplier for the variable step-size BDF(4) method.

Lemma 46 (Multiplier for variable step-size BDF(4))**.** *Let $\delta(z) = \sum_{i=0}^{4} \delta_i z^i$ be the characteristic* BDF(4) *polynomial* (3.6.2)*, where the coefficients δ_i for $i = 0, \dots, 4$, which depend on the step-size ratios κ_4, κ_3 and κ_2, are given in* (3.6.3)*. Assume that the roots z_1, z_2 and z_3 given in* (3.6.4) *of δ lie outside the complex sphere, i.e., $|z_1| \geq 1$, $|z_2| \geq 1$ and $|z_3| \geq 1$. Then, there exists a multiplier $\eta > 0$ such that*

$$\operatorname{Re} \frac{\delta(z)}{1-\eta z} > 0 \qquad \text{for } |z| < 1.$$

Moreover, let $z = x + \mathrm{i}y$ with $x, y \in \mathbb{R}$ and define

$$\mu(z) \coloneqq \mu(z, \kappa_4, \kappa_3, \kappa_2) \coloneqq \prod_{i=0}^{3} (z - z_i(\kappa_4, \kappa_3, \kappa_2)). \tag{3.6.5}$$

If $\operatorname{Re}(\mu(z))x + \operatorname{Im}(\mu(z))y < 0$ holds for all z with $|z| = 1$, then the multiplier satisfies

$$\eta > \max_{|z|=1} \frac{\operatorname{Re}(\mu(z))}{\operatorname{Re}(z)\operatorname{Re}(\mu(z)) + \operatorname{Im}(z)\operatorname{Im}(\mu(z))}. \tag{3.6.6}$$

Proof. Analogous to the variable step-size BDF(3) scheme, we obtain by factoring the polynomial δ in terms of its roots

$$\frac{\delta(z)}{1-\eta z} = \frac{\delta_4(z-1)(z-z_1)(z-z_2)(z-z_3)(1-\eta\overline{z})}{|1-\eta z|^2} = \frac{\delta_4\mu(z)(1-\eta\overline{z})}{|1-\eta z|^2}.$$

Hence, we need to resolve the inequality

$$\operatorname{Re}\{\delta_4\mu(z)(1-\eta\overline{z})\} > 0$$

for $|z| < 1$ or, equivalently, since δ_4 given in (3.6.3) satisfies $\delta_4(\kappa_4, \kappa_3, \kappa_2) > 0$ for $\kappa_4, \kappa_3, \kappa_2 > 0$,

$$\operatorname{Re}\{\mu(z)(1-\eta\overline{z})\} > 0$$

for $|z| < 1$. Thus, with $z = x + \mathrm{i}y$, $x, y \in \mathbb{R}$, we obtain

$$\begin{aligned}\operatorname{Re}\{\mu(z)(1-\eta\overline{z})\} &= \operatorname{Re}(\mu(z))(1-\eta x) - \operatorname{Im}(\mu(z))\eta y \\ &= -\eta(x\operatorname{Re}(\mu(z)) + y\operatorname{Im}(\mu(z))) + \operatorname{Re}(\mu(z))\end{aligned}$$

Considering the edge case $|z| = 1$, we conclude with $\operatorname{Re}(\mu(z))x + \operatorname{Im}(\mu(z))y < 0$

$$\eta > \frac{\operatorname{Re}(\mu(z))}{x \operatorname{Re}(\mu(z)) + y \operatorname{Im}(\mu(z))}$$

if all roots are outside the complex sphere. □

Analogously to the variable step-size BDF(3) scheme, we are only interested in multipliers $\eta < 1$. To determine an upper bound $\kappa^\star$, such that the variable step-size BDF(4) scheme is G-stable for $\kappa_4, \kappa_3, \kappa_2 \in (0, \kappa^\star)$ with a multiplier $\eta < 1$, we define

$$f_0(\kappa) \coloneqq \max_{z \in \mathbb{C} \,:\, |z|=1} \left(\frac{\operatorname{Re}(\mu(z, \kappa, \kappa, \kappa))}{\operatorname{Re}(z) \operatorname{Re}(\mu(z, \kappa, \kappa, \kappa)) + \operatorname{Im}(z) \operatorname{Im}(\mu(z, \kappa, \kappa, \kappa))} \right) - 1 \tag{3.6.7}$$

for $\kappa > 0$, where μ is defined in (3.6.5). Following the idea of Section 3.5, we thus compute the smallest positive root of f_0 using a root-finding algorithm, such as bisection or Brent's method [37, Cpt. 3-4]. We found the smallest positive root $\kappa^\star \approx 1.150$. Further, we checked that the assumptions of Lemma 46 are satisfied for $\kappa_4, \kappa_3, \kappa_2 \in (0, \kappa^\star)$, i.e., that the roots (3.6.4) lie outside the complex sphere and that $\operatorname{Re}(\mu(z))x + \operatorname{Im}(\mu(z))y < 0$. In contrast to the results in [39], the authors obtained zero-stability for the variable step-size BDF(4) scheme for $\kappa_4, \kappa_3, \kappa_2 \in (0, \widetilde{\kappa})$, where $\widetilde{\kappa} \approx 1.101$. We summarize some bounds on the step-ratios for the variable step-size BDF(4) scheme in the following remark.

Remark 47. *We may increase the bound on κ_4, κ_3 and κ_2 to ensure G-stability for the variable step-size* BDF(4) *by freezing the time-step after coarsening. Thus, after switching from a time-step τ to $\kappa\tau$, three steps are taken with this new time-step size $\kappa\tau$. This leads to three scenarios that need to be analyzed: (I) $\tau \to \tau \to \tau \to \kappa\tau$, (II) $\tau \to \tau \to \kappa\tau \to \kappa\tau$ and (III) $\tau \to \kappa\tau \to \kappa\tau \to \kappa\tau$. Further, we consider the case (0) $\tau \to \kappa\tau \to \kappa^2\tau \to \kappa^3\tau$, where we coarsen each step by κ. We obtain the following bounds.*

(0) With the largest positive root $\kappa^\star \approx 1.150$ of f_0 defined in (3.6.7)*, we find that the variable step-size* BDF(4) *scheme is G-stable with a multiplier $\eta < 1$ for*

$$\kappa_4 = \kappa_3 = \kappa_2 = \kappa \quad \textit{with} \quad \kappa \in (0, 1.150).$$

(I) To determine the largest $\kappa_I^\star$ such that $\eta < 1$, we define

$$f_I(\kappa) \coloneqq \max_{z \in \mathbb{C} \,:\, |z|=1} \left(\frac{\operatorname{Re}(\mu(z, \kappa, 1, 1))}{\operatorname{Re}(z) \operatorname{Re}(\mu(z, \kappa, 1, 1)) + \operatorname{Im}(z) \operatorname{Im}(\mu(z, \kappa, 1, 1))} \right) - 1$$

and apply the root-finding algorithm of Brent [37, Cpt. 3-4*] to compute the smallest positive root of f_I, yielding $\kappa_I^\star \approx 1.308$. The variable step-size* BDF(4) *scheme is then G-stable with a multiplier $\eta < 1$ for*

$$\kappa_4 = \kappa, \quad \kappa_3 = \kappa_2 = 1 \quad \textit{with} \quad \kappa \in (0, 1.308).$$

(II) To determine the largest $\kappa_{II}^\star$ such that $\eta < 1$, we define

$$f_{II}(\kappa) \coloneqq \max_{z\in\mathbb{C}\,:\,|z|=1} \left(\frac{\operatorname{Re}(\mu(z,1,\kappa,1))}{\operatorname{Re}(z)\operatorname{Re}(\mu(z,1,\kappa,1)) + \operatorname{Im}(z)\operatorname{Im}(\mu(z,1,\kappa,1))} \right) - 1$$

and apply the root-finding algorithm of Brent [37, Cpt. 3-4] to compute the smallest positive root of f_{II}, yielding $\kappa_{II}^\star \approx 1.367$. The variable step-size BDF(4) *scheme is then G-stable with a multiplier $\eta < 1$ for*

$$\kappa_3 = \kappa, \quad \kappa_4 = \kappa_2 = 1 \quad \textit{with} \quad \kappa \in (0, 1.367).$$

(III) To determine the largest $\kappa_{III}^\star$ such that $\eta < 1$, we define

$$f_{III}(\kappa) \coloneqq \max_{z\in\mathbb{C}\,:\,|z|=1} \left(\frac{\operatorname{Re}(\mu(z,1,1,\kappa))}{\operatorname{Re}(z)\operatorname{Re}(\mu(z,1,1,\kappa)) + \operatorname{Im}(z)\operatorname{Im}(\mu(z,1,1,\kappa))} \right) - 1$$

and apply the root-finding algorithm of Brent [37, Cpt. 3-4] to compute the smallest positive root of f_{III}, yielding $\kappa_{III}^\star \approx 2.792$. The variable step-size BDF(4) *scheme is then G-stable with a multiplier $\eta < 1$ for*

$$\kappa_2 = \kappa, \quad \kappa_4 = \kappa_3 = 1 \quad \textit{with} \quad \kappa \in (0, 2.792).$$

3.7 Discrete orthogonal convolution kernels

In this section, we consider the variable step-size BDF(k) method for $k = 1, 2, 3$ from a different viewpoint. First of all, we start by rewriting the coefficients of the BDF(k) scheme using discrete convolution (DC) kernels. Second, we define the corresponding discrete orthogonal convolution (DOC) kernels. Then, the connection between the DC and the DOC kernels is characterized by the discrete orthogonality identity, which is proven later in Lemma 48. By assuming sufficient bounds on consecutive time-steps, we obtain bounds for the DOC kernels. We require these bounds later on to provide stability bounds for the variable step-size BDF(k) method for the LLG equation. In this section, we follow [39, 92–94].

We consider the variable step-size BDF(k) method, $k = 1, 2, 3$, on the partition $0 = t_0 < t_1 < \cdots < t_n = T$ of the time interval for $n \geq k$ with the time-steps $\tau_i = t_i - t_{i-1}$ for $1 \leq i \leq n$. Further, we incorporate the time-step ratios $\kappa_i = \tau_i/\tau_{i-1}$ for $2 \leq i \leq n$ as defined in (3.1.9). Let $\{u^n\}_{n=0}^N$ be a given sequence. Next, define the rational functions

$$\begin{aligned} b_0(x,y) &\coloneqq \frac{1+2x}{1+x} + \frac{xy}{1+y+xy}, \\ b_1(x,y) &\coloneqq -\frac{x^2}{1+x} - \frac{x^2 y}{1+y+xy} - \frac{x^2y^2}{1+y+xy}\frac{1+x}{1+y}, \\ b_2(x,y) &\coloneqq \frac{x^2y^3}{1+y+xy}\frac{1+x}{1+y} \end{aligned} \tag{3.7.1}$$

for $x, y \geq 0$ and the backward difference $\nabla_\tau u^i \coloneqq \nabla_\tau^1 u^i = u^i - u^{i-1}$ for $1 \leq i \leq n$, where ∇_τ^1 is defined in (3.1.4). Then, the variable step-size BDF(k) scheme for $k = 1, 2, 3$ is given by

$$\tau_n v^n = \sum_{j=1}^{n} b_{n-j}^{(k,n)} \nabla_\tau u^j \tag{3.7.2}$$

for $n \geq k$, where the DC kernels are defined by

$$\begin{aligned} b_0^{(1,n)} &\coloneqq b_0(0,0) = 1, && \text{for } n \geq 1, \\ b_j^{(2,n)} &\coloneqq b_j(\kappa_n, 0), && \text{for } j = 0,1 \text{ and } n \geq 2, \\ b_j^{(3,n)} &\coloneqq b_j(\kappa_n, \kappa_{n-1}), && \text{for } j = 0,1,2 \text{ and } n \geq 3, \end{aligned} \tag{3.7.3}$$

and $b_j^{(k,n)} \coloneqq 0$ for $n \geq j+1 \geq k+1$. Further, the corresponding DOC kernels are recursively defined by

$$\theta_0^{(k,n)} \coloneqq \frac{1}{b_0^{(k,n)}} \quad \text{and} \quad \theta_{n-j}^{(k,n)} \coloneqq -\frac{1}{b_0^{(k,j)}} \sum_{i=j+1}^{n} \theta_{n-i}^{(k,n)} b_{i-j}^{(k,i)} \tag{3.7.4}$$

for $k \leq j \leq n-1$. The following lemma establishes an important property of the DC and DOC kernels, which additionally describes their connection.

Lemma 48 (Discrete orthogonality identity)**.** *Let $\{b_{n-j}^{(k,n)}\}_{j=k}^n$ be the DC kernels defined in* (3.7.3) *and $\{\theta_{n-j}^{(k,n)}\}_{j=k}^n$ be the DOC kernels defined in* (3.7.4)*. Then, the discrete orthogonality identities*

$$\sum_{j=\ell}^{n} \theta_{n-j}^{(k,n)} b_{j-\ell}^{(k,j)} = \delta_{n\ell}, \qquad \sum_{j=\ell}^{n} b_{n-j}^{(k,n)} \theta_{j-\ell}^{(k,j)} = \delta_{n\ell} \tag{3.7.5}$$

hold for $k \leq \ell \leq n$.

Proof. By definition of the DOC kernels (3.7.4), we immediately obtain the first discrete orthogonality identity since $\theta_0^{(k,n)} b_0^{(k,n)} = 1$ and

$$\sum_{j=\ell}^{n} \theta_{n-j}^{(k,n)} b_{j-\ell}^{(k,j)} = \theta_{n-\ell}^{(k,n)} b_0^{(k,\ell)} + \sum_{j=\ell+1}^{n} \theta_{n-j}^{(k,n)} b_{j-\ell}^{(k,j)} = 0$$

for $k \leq \ell \leq n-1$. We prove the second identity by induction. For $\ell = n$, we have again $b_0^{(k,n)} \theta_0^{(k,n)} = 1$. Since $\theta_1^{(k,n)} = -\frac{1}{b_0^{(k,n-1)}} \theta_0^{(k,n)} b_1^{(k,n)}$, we have for $\ell = n-1$

$$\begin{aligned} \sum_{j=n-1}^{n} b_{n-j}^{(k,n)} \theta_{j-\ell}^{(k,j)} &= b_1^{(k,n)} \theta_0^{(k,n-1)} + b_0^{(k,n)} \theta_1^{(k,n)} \\ &= b_1^{(k,n)} \theta_0^{(k,n-1)} - b_0^{(k,n)} \frac{1}{b_0^{(k,n-1)}} \theta_0^{(k,n)} b_1^{(k,n)} \\ &= b_1^{(k,n)} \left(\theta_0^{(k,n-1)} - \frac{1}{b_0^{(k,n-1)}} \right) = 0. \end{aligned}$$

For $k \leq \ell \leq n-2$, we have using the definition of the DOC kernels (3.7.4)

$$\begin{aligned}\sum_{j=\ell}^{n} b_{n-j}^{(k,n)} \theta_{j-\ell}^{(k,j)} &= b_{n-\ell}^{(k,n)} \theta_0^{(k,\ell)} + \sum_{j=\ell+1}^{n} b_{n-j}^{(k,n)} \theta_{j-\ell}^{(k,j)} \\ &= b_{n-\ell}^{(k,n)} \theta_0^{(k,\ell)} - \frac{1}{b_0^{(k,\ell)}} \sum_{j=\ell+1}^{n} b_{n-j}^{(k,n)} \sum_{i=\ell+1}^{j} \theta_{j-i}^{(k,j)} b_{i-\ell}^{(k,i)}.\end{aligned}$$

Subsequently, by changing the order of summation and using induction, we conclude

$$\begin{aligned}\sum_{j=\ell}^{n} b_{n-j}^{(k,n)} \theta_{j-\ell}^{(k,j)} &= b_{n-\ell}^{(k,n)} \theta_0^{(k,\ell)} - \frac{1}{b_0^{(k,\ell)}} \sum_{i=\ell+1}^{n} b_{i-\ell}^{(k,i)} \sum_{j=i}^{n} b_{n-j}^{(k,n)} \theta_{j-i}^{(k,j)} \\ &= b_{n-\ell}^{(k,n)} \theta_0^{(k,\ell)} - \frac{1}{b_0^{(k,\ell)}} \sum_{i=\ell+1}^{n} b_{i-\ell}^{(k,i)} \delta_{ni} = b_{n-\ell}^{(k,n)} \left(\theta_0^{(k,\ell)} - \frac{1}{b_0^{(k,\ell)}} \right) = 0.\end{aligned}$$

□

Utilizing the discrete orthogonality of the kernels, we are able to establish a connection between u^n and $\sum_{j=k}^{n} \tau_j v^j$. This is an important property that we require later in the stability analysis of the BDF scheme for the LLG equation. We start by multiplying (3.7.2) with the DOC kernel $\theta_{\ell-m}^{(k,\ell)}$ and summing over $m = k$ to ℓ

$$\sum_{m=k}^{\ell} \theta_{\ell-m}^{(k,\ell)} \tau_m v^m = \sum_{m=k}^{\ell} \theta_{\ell-m}^{(k,\ell)} \sum_{j=1}^{m} b_{m-j}^{(k,m)} \nabla_\tau u^j.$$

Subsequently, by separating the initial steps, changing the order of summation, and employing the discrete orthogonality identity (3.7.5), we obtain

$$\begin{aligned}\sum_{m=k}^{\ell} \theta_{\ell-m}^{(k,\ell)} \tau_m v^m &= \sum_{j=1}^{k-1} \nabla_\tau u^j \sum_{m=k}^{\ell} \theta_{\ell-m}^{(k,\ell)} b_{m-j}^{(k,m)} + \sum_{j=k}^{\ell} \nabla_\tau u^j \sum_{m=j}^{\ell} \theta_{\ell-m}^{(k,\ell)} b_{m-j}^{(k,m)} \\ &= \mathcal{I}_k^m + \nabla_\tau u^\ell\end{aligned}$$

for $\ell \geq k$, where we defined

$$\mathcal{I}_k^\ell := \sum_{j=1}^{k-1} \nabla_\tau u^j \sum_{m=k}^{\ell} \theta_{\ell-m}^{(k,\ell)} b_{m-j}^{(k,m)}.$$

Finally, summing over $\ell = k$ to n and exchanging the order of summation again yields

$$u^n - u^{k-1} + \sum_{\ell=k}^{n} \mathcal{I}_k^\ell = \sum_{\ell=k}^{n} \sum_{m=k}^{\ell} \theta_{\ell-m}^{(k,\ell)} \tau_m v^m = \sum_{m=k}^{n} \left(\tau_m v^m \sum_{\ell=m}^{n} \theta_{\ell-m}^{(k,\ell)} \right). \tag{3.7.6}$$

Remark 49. *In [92, 94], the functions characterizing the BDF scheme are defined slightly*

differently. They define the rational functions

$$d_0(x,y) \coloneqq b_0(x,y), \qquad d_1(x,y) \coloneqq \frac{1}{x} b_1(x,y), \qquad d_2(x,y) \coloneqq \frac{1}{xy} b_1(x,y).$$

Then, the backward differentiation formula are defined by

$$v^n \coloneqq \sum_{j=1}^{n} d_{n-j}^{(k,n)} \frac{\nabla_\tau u^j}{\tau_j},$$

where $d_{n-j}^{(k,n)}$ *is defined analogously to* (3.7.3)*. The corresponding DOC kernels are recursively defined by*

$$\vartheta_0^{(k,n)} \coloneqq \frac{1}{d_0^{(k,n)}}, \qquad \vartheta_{n-j}^{(k,n)} \coloneqq -\frac{1}{d_0^{(k,j)}} \sum_{i=j+1}^{n} \vartheta_{n-i}^{(k,n)} d_{i-j}^{(k,i)}, \tag{3.7.7}$$

for $k \le j \le n-1$*.*

The connection between the DOC kernels $\{\theta_{n-j}^{(k,n)}\}_{j=1}^n$ and $\{\vartheta_{n-j}^{(k,n)}\}_{j=1}^n$ is given by the following lemma.

Lemma 50. *Let* $\{\theta_{n-j}^{(k,n)}\}_{j=k}^n$ *be defined by* (3.7.4) *and* $\{\vartheta_{n-j}^{(k,n)}\}_{j=k}^n$ *be defined by* (3.7.7)*. Then, for* $3 \le j \le n$*, the DOC kernels satisfy*

$$\theta_{n-j}^{(k,n)} = \left(\prod_{i=j+1}^{n} \kappa_i \right) \vartheta_{n-j}^{(k,n)}.$$

Proof. For readability, we abbreviate $\theta_{n-j}^{(n)} = \theta_{n-j}^{(k,n)}$ and $\vartheta_{n-j}^{(n)} = \vartheta_{n-j}^{(k,n)}$. Let $n \ge 3$. Then we have for $j = n$ that

$$\theta_0^{(n)} = \frac{1}{b_0^{(n)}} = \frac{1}{d_0^{(n)}} = \vartheta_0^{(n)}.$$

For $j = n-1$ we have

$$\theta_1^{(n)} = -\frac{1}{b_0^{(n-1)}} \theta_0^{(n)} b_1^{(n)} = -\frac{1}{d_0^{(n-1)}} \vartheta_0^{(n)} \kappa_n d_1^{(n)} = \kappa_n \vartheta_1^{(n)}.$$

By induction, we obtain for $0 \le j \le n-2$ that

$$\begin{aligned}
\theta_{n-j}^{(n)} &= -\frac{1}{b_0^{(j)}} \left(\theta_{n-j-1}^{(n)} b_1^{(j+1)} + \theta_{n-j-2}^{(n)} b_2^{(j+2)} \right) \\
&= -\frac{1}{d_0^{(j)}} \left(\theta_{n-j-1}^{(n)} \kappa_{j+1} d_1^{(j+1)} + \theta_{n-j-2}^{(n)} \kappa_{j+2} \kappa_{j+1} d_2^{(j+2)} \right) \\
&= -\frac{1}{d_0^{(j)}} \left(\left(\prod_{i=j+2}^{n} \kappa_i \right) \vartheta_{n-j-1}^{(n)} \kappa_{j+1} d_1^{(j+1)} + \left(\prod_{i=j+3}^{n} \kappa_i \right) \vartheta_{n-j-2}^{(n)} \kappa_{j+2} \kappa_{j+1} d_2^{(j+2)} \right)
\end{aligned}$$

$$= \left(\prod_{i=j+1}^{n} \kappa_i \right) \vartheta_{n-j}^{(n)}.$$

□

An explicit formula for the DOC kernels for the variable step-size BDF(2) method is presented in the following lemma, which is provided in [93, Lemma 2.3].

Lemma 51 (Explicit formula for BDF(2))**.** *The DOC kernels* $\{\theta_{n-j}^{(2,n)}\}_{j=2}^{n}$ *defined in* (3.7.4) *satisfy the explicit formula*

$$\theta_{n-j}^{(2,n)} = \frac{1}{b_0^{(2,j)}} \prod_{i=j+1}^{n} \frac{\kappa_i^2}{1+2\kappa_i} = \frac{1+\kappa_j}{1+2\kappa_j} \prod_{i=j+1}^{n} \frac{\kappa_i^2}{1+2\kappa_i} \tag{3.7.8}$$

for $2 \le j \le n$.

Proof. By the definition of the convolution kernels (3.7.3), we have

$$\frac{b_1^{(2,n)}}{b_0^{(2,n)}} = -\frac{\kappa_n}{1+2\kappa_n}$$

for $n \ge 2$. Next, using the definition of the DOC kernels (3.7.4) and the above equation, we obtain

$$\theta_{n-j}^{(2,n)} b_0^{(2,j)} = -\theta_{n-j-1}^{(2,n)} b_1^{(2,j+1)} = \frac{\kappa_{j+1}}{1+2\kappa_{j+1}} \theta_{n-j-1}^{(2,n)} b_0^{(2,j+1)}$$

for $2 \le j \le n-1$. Since $\theta_0^{(2,n)} b_0^{(2,n)} = 1$, by recursively inserting the above equation into itself, we obtain

$$\theta_{n-j}^{(2,n)} b_0^{(2,j)} = \prod_{i=j+1}^{n} \frac{\kappa_i^2}{1+2\kappa_i}.$$

This concludes the proof. □

To obtain a bound for the DOC kernels $\{\theta_{n-j}^{(2,n)}\}_{j=2}^{n}$, we can utilize the explicit formula as demonstrated later in the proof of Theorem 57. In contrast, no explicit formula is known for the DOC kernels $\{\theta_{n-j}^{(3,n)}\}_{j=3}^{n}$ of the BDF(3) scheme. Thus, we must employ a different strategy to obtain an upper bound. We follow the technique in [39], where Calvo et al. proved the zero-stability of variable step-size BDF formulas of up to order 5, which has been adapted in [92] to the DOC kernels.

With the definition (3.7.1) of the rational functions b_0, b_1 and b_2, we define

$$\alpha(x,y) := -\frac{b_1(x,y)}{b_0(x,y)} = \frac{x^2\big((1+y+xy)(1+y)+y(1+x)(1+y)+y^2(1+x)^2\big)}{(1+y)\big((1+2x)(1+y+xy)+xy(1+x)\big)} \tag{3.7.9}$$

$$= \frac{x^2(x^2y^2+4xy^2+3y^2+2xy+3y+1)}{(1+y)(3x^2y+4xy+2x+y+1)} \tag{3.7.10}$$

and

$$\beta(x,y) \coloneqq \frac{b_2(x,y)}{b_0(x,y)} = \frac{x^2y^3(1+x)^2}{(1+y)(3x^2y+4xy+2x+y+1)}.$$

for $x, y > 0$. Next, we define the coefficients

$$\alpha_m \coloneqq -\frac{b_1^{(3,m)}}{b_0^{(3,m)}} = \alpha(\kappa_m, \kappa_{m-1}), \qquad \beta_m \coloneqq \frac{b_2^{(3,m)}}{b_0^{(3,m)}} = \beta(\kappa_m, \kappa_{m-1}) \tag{3.7.11}$$

for $m \geq 3$ and the scaled DOC kernels

$$\begin{aligned} \widehat{\theta}_{m-\ell}^{(3,m)} &\coloneqq \theta_{m-\ell}^{(3,m)} b_0^{(3,\ell)} && \text{for } m \geq \ell \geq 3, \\ \widehat{\boldsymbol{\theta}}_{m-\ell}^{(3,m)} &\coloneqq \begin{bmatrix} \widehat{\theta}_{m-\ell}^{(3,m)} \\ \widehat{\theta}_{m-1-\ell}^{(3,m-1)} \end{bmatrix} && \text{for } m \geq \ell+1 \geq 4. \end{aligned} \tag{3.7.12}$$

Then, the second discrete orthogonality identity in (3.7.5) combined with the scaled DOC kernels defined in (3.7.12) yields

$$\widehat{\theta}_{m-\ell}^{(3,m)} b_0^{(3,m)} + \widehat{\theta}_{m-1-\ell}^{(3,m-1)} b_1^{(3,m)} + \widehat{\theta}_{m-2-\ell}^{(3,m-2)} b_2^{(3,m)} = \delta_{m\ell} b_0^{(3,m)}.$$

Thus, with the definitions of α_m and β_m given in (3.7.11), we obtain the two-step recursion

$$\widehat{\theta}_{m-\ell}^{(3,m)} - \alpha_m \widehat{\theta}_{m-1-\ell}^{(3,m-1)} + \beta_m \widehat{\theta}_{m-2-\ell}^{(3,m-2)} = 0 \tag{3.7.13}$$

for $m \geq \ell + 2 \geq 5$. Let μ be a complex constant with $\operatorname{Im}(\mu) \neq 0$. Then, following [39], we define for $A \in \mathbb{C}^{2\times 2}$ the norm

$$\|A\|_{\mu,\infty} \coloneqq \|\mathcal{M}_\mu^{-1} A \mathcal{M}_\mu\|_\infty, \qquad \mathcal{M}_\mu \coloneqq \begin{bmatrix} \mu & \overline{\mu} \\ 1 & 1 \end{bmatrix}, \tag{3.7.14}$$

where the $\|\cdot\|_\infty$ is defined in (1.4.2). Defining the matrix

$$A_m \coloneqq \begin{bmatrix} \alpha_m & -\beta_m \\ 1 & 0 \end{bmatrix} \tag{3.7.15}$$

for $m \geq 3$, we rewrite the two-step recursion (3.7.13) to

$$\widehat{\boldsymbol{\theta}}_{m-\ell}^{(3,m)} = A_m \widehat{\boldsymbol{\theta}}_{m-1-\ell}^{(3,m-1)} \tag{3.7.16}$$

for $m \geq \ell + 2 \geq 5$. Inserting (3.7.16) recursively into itself, we obtain

$$\widehat{\boldsymbol{\theta}}_{n-\ell}^{(3,n)} = \prod_{m=\ell+2}^{n} A_m \widehat{\boldsymbol{\theta}}_1^{(3,\ell+1)}$$

for $n \geq \ell + 2 \geq 5$. Finally, by multiplying each A_m for $m \in \{\ell+2, \ldots, n\}$ with $\mathcal{M}_\mu \mathcal{M}_\mu^{-1} = I_2$

from both sides, we estimate

$$\|\widehat{\boldsymbol{\theta}}_{n-\ell}^{(3,n)}\|_\infty \le \|\mathcal{M}_\mu\|_\infty \prod_{m=\ell+2}^{n} \|A_m\|_{\mu,\infty}\|\mathcal{M}_\mu^{-1}\|_\infty \|\widehat{\boldsymbol{\theta}}_1^{(3,\ell+1)}\|_\infty \tag{3.7.17}$$

for $n \ge \ell + 2 \ge 5$, where the ℓ^∞-norm is defined in (1.4.1). The optimal value $\kappa^\star \approx 1.476$, such that $\|A_m\|_{H,\infty} \le 1$ for $\kappa_m, \kappa_{m-1} \in (0, \kappa^\star)$, has been computed numerically in [39] and yields the tangential point $\mu^\star \approx 0.4943 + 0.5748\,\mathrm{i}$. This value coincides with the G-stability bound we derived earlier in Section 3.5 for the variable step-size BDF(3) method, as well as with the zero-stability bound given in [39]. With the above calculations at hand, we obtain the following lemma, which provides a bound for the DOC kernels of the variable step-size BDF(3) method and is adapted from [92, Lemma 3.1].

Lemma 52. *If the step ratios $\kappa_j \le \kappa^\star \approx 1.476$ for $j \ge 2$, there exists a positive constant C_3 such that the DOC kernels $\{\theta_{n-j}^{(3,n)}\}_{j=3}^n$ defined in* (3.7.4) *satisfy*

$$\sum_{j=3}^{n} |\theta_{n-j}^{(3,n)}| \le C_3, \qquad \sum_{j=\ell}^{n} |\theta_{j-\ell}^{(3,j)}| \le C_3 \tag{3.7.18}$$

for $n \ge 3$ and $\ell \ge 3$, where C_3 is independent of the time t_n and the step ratios κ_j.

Proof. Since $\|A_m\|_{\mu,\infty}$ is a continuous function w.r.t. $\kappa_m, \kappa_{m-1} \in (0, \kappa^\star)$ for A_m defined in (3.7.15), we can find $\delta > 0$, such that

$$\|A_m\|_{\mu,\infty} \le \delta < 1$$

for $m \ge 3$. By the definition of b_0 in (3.7.3), we have

$$b_0^{(3,m)} \ge 1 \qquad \text{for } \kappa_m, \kappa_{m-1} \ge 0 \text{ and } m \ge 3. \tag{3.7.19}$$

Thus, with the definition of the DOC kernels (3.7.4) and (3.7.19), we obtain

$$\begin{aligned} \left|\theta_0^{(3,m)}\right| &= \left|\frac{1}{b_0^{(3,m)}}\right| \le 1 \le \frac{1}{\delta}, \\ \left|\theta_1^{(3,m+1)}\right| &= \left|\frac{b_1^{(3,m+1)}}{b_0^{(3,m)}}\theta_0^{(3,m+1)}\right| \le \left|\frac{b_1^{(3,m+1)}}{b_0^{(3,m+1)}}\right| = \alpha_{m+1} \end{aligned}$$

for $m \ge 3$. Further, by the definition of the scaled DOC kernels (3.7.12), we have

$$\widehat{\theta}_1^{(3,m+1)} = \frac{b_1^{(3,m+1)}}{b_0^{(3,m+1)}} = \alpha_{m+1}, \qquad \widehat{\theta}_0^{(3,m)} = 1,$$

which leads to

$$\|\widehat{\boldsymbol{\theta}}_1^{(3,m+1)}\|_\infty = \max\{\alpha_{m+1}, 1\} \tag{3.7.20}$$

for $m \geq 3$. Next, applying (3.7.17), (3.7.19) and (3.7.20), we obtain

$$\left|\theta_{n-\ell}^{(3,n)}\right| \leq \left|\widehat{\theta}_{n-\ell}^{(3,n)}\right| \leq \|\widehat{\boldsymbol{\theta}}_{n-\ell}^{(3,n)}\|_\infty \leq \|\mathcal{M}_\mu\|_\infty \|\mathcal{M}_\mu^{-1}\|_\infty \max\{\alpha_{\ell+1}, 1\}\delta^{n-\ell-1}$$

for $n \geq \ell + 2 \geq 5$. Defining

$$c_{\mu,\ell} \coloneqq \max\{\|\mathcal{M}_\mu\|_\infty \|\mathcal{M}_\mu^{-1}\|_\infty, 1\} \max\{\alpha_\ell, 1\}, \qquad c_\mu = \max_{j=\ell,\dots,n} \{c_{\mu,j}\} \tag{3.7.21}$$

for $n \geq \ell \geq 3$ and combining the above inequality yields

$$\sum_{j=\ell}^{n} |\theta_{j-\ell}^{(3,j)}| \leq c_\mu \sum_{j=\ell}^{n} \delta^{j-\ell-1} < \frac{c_\mu}{\delta(1-\delta)} =: C_3$$

for $n \geq \ell \geq 3$, where we estimated the partial geometric series by its limiting value in the last step. Similarly, we obtain, for $n \geq 3$,

$$\sum_{j=3}^{n} |\theta_{n-j}^{(3,n)}| \leq c_\mu \sum_{j=3}^{n} \delta^{n-j-1} < \frac{c_\mu}{\delta(1-\delta)} = C_3.$$

□

Next, we want to make some comments about the constant (3.7.21) that appeared in the previous proofs. Taking $\mu = \frac{1}{2} + \frac{1}{2}\,\mathrm{i}$, which is close to the optimal $\mu^\star \approx 0.4943 + 0.5748\,\mathrm{i}$, we obtain $\|\mathcal{M}_\mu\|_\infty = 2$ and $\|\mathcal{M}_\mu^{-1}\|_\infty = 1 + \frac{1}{\sqrt{2}}$. Further, since α defined in (3.7.9) is monotonically increasing for $x, y \geq 0$, cf. Appendix A.2, we obtain

$$\alpha_\ell \leq \alpha(\kappa^\star, \kappa^\star) \approx 1.255$$

for $\ell \geq 3$. Combined, we obtain

$$c_\mu \leq (2 + \sqrt{2})\alpha(\kappa^\star, \kappa^\star) \approx 4.285\,.$$

Remark 53. *To compute the optimal $\kappa^\star$ for A_m defined in* (3.7.15)*, we have to check that*

$$\|A_m(\kappa_m, \kappa_{m-1})\|_{\mu,\infty} \leq \delta < 1$$

for all $\kappa_m, \kappa_{m-1} \in (0, \kappa^\star)$. Explicitly computing the matrix products in (3.7.14) *for A_m yields*

$$\|A_m\|_{\mu,\infty} = \frac{1}{|\mu - \overline{\mu}|} \max\Big\{|-\overline{\mu}^2 + \alpha_m - \beta_m| + |-|\mu|^2 + \alpha_m\mu - \beta_m|,$$
$$|\mu^2 - \alpha_m + \beta_m| + ||\mu|^2 - \alpha_m\overline{\mu} + \beta_m|\Big\}$$

Noticing that both arguments in the maximum are equal leads to

$$\|A_m\|_{\mu,\infty} = \frac{|\mu^2 - \alpha_m\mu + \beta_m| + ||\mu|^2 - \alpha_m\overline{\mu} + \beta_m|}{|\mu - \overline{\mu}|}. \tag{3.7.22}$$

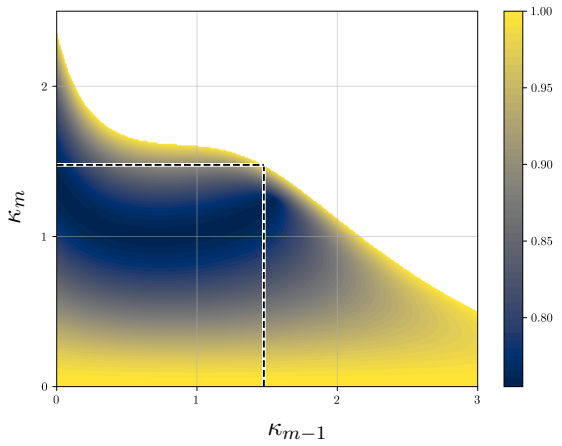


Figure 3.7.1: Values of $\|A_m\|_{\mu^\star,\infty}$ for $\mu^\star \approx 0.4943 + 0.5748\,\mathrm{i}$, if $\|A_m\|_{\mu^\star,\infty} < 1$. In other words, the colored region presents the step ratios where the DOC kernels are bounded according to Lemma 52. The square $(0, \kappa^\star)^2$, where $\kappa^\star \approx 1.476$, marked with dashed lines is completely contained in the colored region.

By direct calculations, we obtain

$$\left||\mu|^2 - \alpha_m\overline{\mu} + \beta_m\right|^2 = \left||\mu|^2 - \alpha_m\mu + \beta_m\right|^2 = \left|\mu^2 - \alpha_m\mu + \beta_m\right|^2 + \beta_m\,|\mu - \overline{\mu}|^2\,.$$

Next, define $g^m(\mu) \coloneqq \frac{|\mu^2 - \alpha_m\mu + \beta_m|}{|\mu - \overline{\mu}|}$ *for* $\mu \in \mathbb{C}$. *Finally, rewriting* (3.7.22) *with the above equations yields*

$$\|A_m\|_{\mu,\infty} = g^m(\mu) + \sqrt{\beta_m + g^m(\mu)^2}.$$

In Figure 3.7.1, we present the values of $\|A_m\|_{\mu^\star,\infty}$ *for* $\mu^\star \approx 0.4943 + 0.5748\,\mathrm{i}$. *Obviously,* $(0, \kappa^\star)^2$, *where* $\kappa^\star \approx 1.476$, *marked with dashed lines, is contained in the colored region. Thus, the DOC kernels* $\{\theta_{n-j}^{(3,n)}\}_{j=3}^{n}$ *are bounded for* $\kappa_n, \kappa_{n-1} \in (0, \kappa^\star)$.

Remark 54. *For the DOC kernels of the variable step-size* BDF(4) *and* BDF(5) *scheme, no bound similar to* (3.7.18) *is known. However, following the numerical computations of [39, Sct. 4], it is reasonable to assume that a bound on the DOC kernels of the variable step-size* BDF(4) *scheme can be obtained by extending* $\mathcal{M}_\mu$ *in* (3.7.14) *to*

$$\mathcal{M}_\mu^{(3)} \coloneqq \begin{bmatrix} \mu & \overline{\mu} & \lambda\delta \\ 1 & 1 & \lambda\gamma \\ \beta & \beta & \lambda \end{bmatrix},$$

where $\mu \in \mathbb{C}$, $\delta, \lambda, \gamma, \beta \in \mathbb{R}$ *have to be chosen such that* $\mathcal{M}_\mu^{(3)}$ *is regular, and following the ideas of Lemma 52.*

3.8 Stability bounds for the VS-BDF method for the LLG equation

In this section, we utilize the previously introduced Nevanlinna–Odeh multipliers to obtain a discrete energy inequality for the LLG equation. Furthermore, by applying the theory of the DOC kernels from Section 3.7, we are able to extend the discrete energy inequality into an $\mathcal{L}^2$-bound for the discrete solution. To ensure the validity of these inequalities, we require an upper bound on the consecutive time-step ratios (3.1.9), as well as a lower bound on the damping coefficient α. In this section, we consider a non-changing spatial mesh with the finite element spaces $V_h = V_h^n = W_h^n$ for all $n \geq 0$, and thus we omit the superscript for the discrete tangent space, writing simply $\boldsymbol{T}_h^n = \boldsymbol{T}_h$. For the convenience of the reader, we recall the fully discrete scheme with the BDF(k), $k \in \{1, \dots, 5\}$, scheme from Section 1.6: Find $\boldsymbol{m}_h^n \in \boldsymbol{V}_h$ with $\boldsymbol{v}_h^n \in \boldsymbol{T}_h(\widehat{\boldsymbol{m}}_h^n)$ such that

$$\alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \ell_{\mathrm{ex}}^2\langle \boldsymbol{\nabla}\boldsymbol{m}_h^n, \boldsymbol{\nabla}\boldsymbol{\varphi}_h^n\rangle = \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n\rangle, \tag{3.8.1}$$

for all $\boldsymbol{\varphi}_h^n \in \boldsymbol{T}_h(\widehat{\boldsymbol{m}}_h^n)$ for $n \geq k$. Further, let $\boldsymbol{q}$ be the unique interpolation polynomial of $\boldsymbol{m}_h^{n-k}, \dots, \boldsymbol{m}_h^{n-1}$ in $\mathbb{P}_{k-1}$ as described in Section 1.6. Then, we consider the predictor

$$\widehat{\boldsymbol{m}}_h^n = \frac{\boldsymbol{q}(t_n)}{|\boldsymbol{q}(t_n)|}$$

for $n \geq k$. Next, to obtain the discrete energy inequality, we require the $\mathcal{L}^2(\Omega)$-orthogonal projection onto the tangent space $\boldsymbol{T}_h(\boldsymbol{m})$, denoted by

$$\boldsymbol{P}_h(\boldsymbol{m}) : \boldsymbol{V}_h \to \boldsymbol{T}_h(\boldsymbol{m}). \tag{3.8.2}$$

Although we do not have an explicit expression for this projection, several properties are stated in [7, Sct. 5]. In particular, we require a Lipschitz-type estimate as provided in [7, Lemma 5.2], which we recall below.

Lemma 55 (Lipschitz-type bound for $\boldsymbol{P}_h$)**.** *Let $\boldsymbol{m} \in \boldsymbol{\mathcal{W}}^{1,\infty}(\Omega)$ and $\widetilde{\boldsymbol{m}} \in \boldsymbol{\mathcal{H}}^1(\Omega)$ with $|\boldsymbol{m}| = |\widetilde{\boldsymbol{m}}| = 1$ a.e. and $\|\boldsymbol{m}\|_{\boldsymbol{\mathcal{W}}^{1,\infty}} \leq R$. Then, there exists a constant $c_R > 0$ and a mesh-size $h_R > 0$, such that*

$$\|(\boldsymbol{P}_h(\boldsymbol{m}) - \boldsymbol{P}_h(\widetilde{\boldsymbol{m}}))\boldsymbol{v}_h\| \leq c_R\|\boldsymbol{m} - \widetilde{\boldsymbol{m}}\|_{\boldsymbol{\mathcal{L}}^\infty}\|\boldsymbol{v}_h\|$$

for $h \leq h_R$ and $\boldsymbol{v}_h \in \boldsymbol{V}_h$.

In [7, Prop. 3.1], a discrete inequality for the fully discrete scheme (3.8.1) with the uniform step-size BDF(2) method is provided. This was extended in our work [36] to variable step-size BDF(k) with $k = 1, 2$. We extend the same type of bound to the variable step-size BDF(k) method for $k = 3, \dots, 5$, provided that the ratio of consecutive time-steps is limited and the damping coefficient α has a positive lower bound based on [36]. Moreover, starting from the

discrete energy inequality we derive an $\mathcal{L}^2$-bound for the discrete solution using the DOC kernels introduced in Section 3.7. In particular, the discrete energy inequality is valid for the uniform step-size BDF(k) method for $k = 3, 4, 5$, where the Nevanlinna–Odeh multipliers are given in Lemma 42. The precise conditions on the step-ratios for $k = 2, 3, 4$ are provided in Section 3.4, 3.5 and 3.6. In principle, it is reasonable to assume that the argument for $k = 3, 4$ can be extended to $k = 5$. However, establishing the precise conditions for the variable-step BDF(5) method requires substantially more work, as computing the roots is even more involved.

Theorem 56 (Energy inequality for orders $k = 1, \dots, 4$)**.** *Consider the discretization* (3.8.1) *of the LLG equation* (1.2.16) *for $k \in \{1, \dots, 4\}$ with finite elements of polynomial degree $p \geq 1$. Further, we assume $\tau = \mathcal{O}(h)$ and, for $k \geq 2$, that the time-steps ratios satisfy $\kappa_n = \tau_n/\tau_{n-1} \leq \kappa_0 < \kappa_{0,k}$ for all $n \geq 2$, where*

$$\kappa_{0,2} = 1 + \sqrt{2}, \qquad \kappa_{0,3} \approx 1.476, \qquad \kappa_{0,4} \approx 1.150$$

and $\alpha \geq \alpha_{0,k} > 0$ for some $\alpha_{0,k}$ that depends on $\kappa_{0,k}$. Then, the numerical solution satisfies the following discrete energy inequality: For $n > k$ it holds

$$\|\boldsymbol{\nabla} \boldsymbol{m}_h^n\|^2 + \alpha \sum_{j=k}^{n} \tau_j \|\boldsymbol{v}_h^j\|^2 \leq C \left(\sum_{i=0}^{k-1} \|\boldsymbol{\nabla} \boldsymbol{m}_h^i\|^2 + \sum_{j=k}^{n} \tau_j \|\boldsymbol{f}_h^j\|^2 \right)$$

under the regularity assumptions (1.6.9)*, where the constant $C > 0$ is independent of the spatial mesh size h and the time-steps τ_j, but depends exponentially on T.*

Proof. We start by recalling that the BDF(k) scheme provides coefficients δ_j, $j = 0, \dots, k$, such that

$$\tau_n \boldsymbol{v}_h^n = \sum_{j=0}^{k} \delta_j \boldsymbol{m}_h^{n-j} \tag{3.8.3}$$

for $n \geq k$, see (1.6.5). According to Sections 3.4, 3.5 and 3.6, when the time-step ratios are bounded, there exists a multiplier $\eta_k \in [0, 1)$ such that

$$\operatorname{Re} \frac{\delta(z)}{1 - \eta_k z} > 0 \qquad \text{for } |z| < 1.$$

The precise conditions on the time-step ratios are given in Lemma 43, Remark 45 and Remark 47. Thus, by Lemma 40, there exists a symmetric positive definite matrix $G = (g_{ij}) \in \mathbb{R}^{k \times k}$, such that

$$\begin{aligned} \langle \sum_{j=0}^{k} \delta_j \boldsymbol{\nabla} \boldsymbol{m}_h^{n-j}, \boldsymbol{\nabla}(\boldsymbol{m}_h^n - \eta \boldsymbol{m}_h^{n-1}) \rangle \geq & \sum_{i,j=1}^{k} g_{ij} \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-i+1}, \boldsymbol{\nabla} \boldsymbol{m}_h^{n-j+1} \rangle \\ & - \sum_{i,j=1}^{k} g_{ij} \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-i}, \boldsymbol{\nabla} \boldsymbol{m}_h^{n-j} \rangle \end{aligned} \tag{3.8.4}$$

for $n \geq k$. As in [7], the main idea is to write down the weak equation for $\boldsymbol{m}_h^n - \eta \boldsymbol{m}_h^{n-1}$ and test it with $\boldsymbol{v}_h^n$. The main point in the proof is to obtain a lower bound for $(\boldsymbol{\nabla}(\boldsymbol{m}_h^n - \eta \boldsymbol{m}_h^{n-1}), \boldsymbol{\nabla} \boldsymbol{v}_h^n)$. Combining (3.8.3), the lower bound (3.8.4) and defining the norm

$$\|\boldsymbol{M}_h^n\|_G^2 \coloneqq \sum_{i,j=1}^{k} g_{ij} \langle \boldsymbol{m}_h^{n-i+1}, \boldsymbol{m}_h^{n-j+1} \rangle$$

with $\boldsymbol{M}_h^n = [\boldsymbol{m}_h^n, \ldots, \boldsymbol{m}_h^{n-k+1}]$ for $n \geq k-1$ and some symmetric positive definite matrix $\boldsymbol{G} \in \mathbb{R}^{k \times k}$, we obtain the lower bound

$$\langle \tau_n \boldsymbol{\nabla} \boldsymbol{v}_h^n, \boldsymbol{\nabla}(\boldsymbol{m}_h^n - \eta \boldsymbol{m}_h^{n-1}) \rangle \geq \|\boldsymbol{M}_h^n\|_G^2 - \|\boldsymbol{M}_h^{n-1}\|_G^2 \tag{3.8.5}$$

for $n \geq k$.

We start, for simplicity, by assuming $\boldsymbol{f}_h^n = \boldsymbol{0}$ for $n \geq k$. Next, we recall the weak equation for $\boldsymbol{v}_h^n$ and $\boldsymbol{v}_h^{n-1}$

$$\begin{aligned} \alpha \langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^n, \boldsymbol{\nabla} \boldsymbol{\varphi}_h^n \rangle = 0, \quad \forall \boldsymbol{\varphi}_h^n \in \boldsymbol{T}_h(\widehat{\boldsymbol{m}}_h^n), \\ \alpha \langle \boldsymbol{v}_h^{n-1}, \boldsymbol{\varphi}_h^{n-1} \rangle + \langle \widehat{\boldsymbol{m}}_h^{n-1} \times \boldsymbol{v}_h^{n-1}, \boldsymbol{\varphi}_h^{n-1} \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-1}, \boldsymbol{\nabla} \boldsymbol{\varphi}_h^{n-1} \rangle = 0, \quad \forall \boldsymbol{\varphi}_h \in \boldsymbol{T}_h(\widehat{\boldsymbol{m}}_h^{n-1}), \end{aligned}$$

for $n > k$. In order to derive an equation for $\boldsymbol{v}_h^n - \eta \boldsymbol{v}_h^{n-1}$, we use the test function $\boldsymbol{v}_h^n \in \boldsymbol{T}_h(\widehat{\boldsymbol{m}}_h^n)$ in the first equation and $\boldsymbol{P}_h^{n-1} \boldsymbol{v}_h^n$ with $\boldsymbol{P}_h^{n-1} \coloneqq \boldsymbol{P}_h(\widehat{\boldsymbol{m}}_h^{n-1}) \in \boldsymbol{T}_h(\widehat{\boldsymbol{m}}_h^{n-1})$ in the second equation, where $\boldsymbol{P}_h(\boldsymbol{m})$ denotes the orthogonal projection onto $\boldsymbol{T}_h(\boldsymbol{m})$ defined in (3.8.2). Rewriting $\boldsymbol{P}_h^{n-1} \boldsymbol{v}_h^n = \boldsymbol{v}_h^n - (\boldsymbol{P}_h^n - \boldsymbol{P}_h^{n-1}) \boldsymbol{v}_h^n = \boldsymbol{v}_h^n - \boldsymbol{p}_h^n$, where $\boldsymbol{p}_h^n \coloneqq (\boldsymbol{P}_h^n - \boldsymbol{P}_h^{n-1}) \boldsymbol{v}_h^n$, we obtain by inserting these test functions

$$\begin{aligned} &\alpha \langle \boldsymbol{v}_h^n, \boldsymbol{v}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^n, \boldsymbol{\nabla} \boldsymbol{v}_h^n \rangle = 0, \\ &\alpha \langle \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-1}, \boldsymbol{\nabla} \boldsymbol{v}_h^n \rangle \\ &\qquad = \alpha \langle \boldsymbol{v}_h^{n-1}, \boldsymbol{p}_h^n \rangle - \langle \widehat{\boldsymbol{m}}_h^{n-1} \times \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n - \boldsymbol{p}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-1}, \boldsymbol{\nabla} \boldsymbol{p}_h^n \rangle \\ &\qquad = \alpha \langle \boldsymbol{v}_h^{n-1}, \boldsymbol{p}_h^n \rangle - \langle \widehat{\boldsymbol{m}}_h^{n-1} \times \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n - \sigma \boldsymbol{v}_h^{n-1} - \boldsymbol{p}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-1}, \boldsymbol{\nabla} \boldsymbol{p}_h^n \rangle \end{aligned}$$

for $n > k$, where we can choose either $\sigma = 0$ or $\sigma = 1$. By subtraction, we obtain

$$\begin{aligned} &\alpha \langle \boldsymbol{v}_h^n - \eta \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla}(\boldsymbol{m}_h^n - \eta \boldsymbol{m}_h^{n-1}), \boldsymbol{\nabla} \boldsymbol{v}_h^n \rangle \\ &\quad = -\eta \alpha \langle \boldsymbol{v}_h^{n-1}, \boldsymbol{p}_h^n \rangle + \eta \langle \widehat{\boldsymbol{m}}_h^{n-1} \times \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n - \sigma \boldsymbol{v}_h^{n-1} - \boldsymbol{p}_h^n \rangle - \eta \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{n-1}, \boldsymbol{\nabla} \boldsymbol{p}_h^n \rangle \end{aligned} \tag{3.8.6}$$

for $n > k$. Utilizing (3.8.5) and $\langle \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n \rangle \leq \frac{1}{2}(\|\boldsymbol{v}_h^{n-1}\|^2 + \|\boldsymbol{v}_h^n\|^2)$, we estimate the left-hand side of (3.8.6) by

$$\begin{aligned} &\alpha \langle \boldsymbol{v}_h^n - \eta \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla}(\boldsymbol{m}_h^n - \eta \boldsymbol{m}_h^{n-1}), \boldsymbol{\nabla} \boldsymbol{v}_h^n \rangle \\ &\qquad \geq (1 - \frac{\eta}{2}) \alpha \|\boldsymbol{v}_h^n\|^2 - \frac{\eta}{2} \alpha \|\boldsymbol{v}_h^{n-1}\|^2 + \frac{\ell_{\mathrm{ex}}^2}{\tau_n} (\|\boldsymbol{\nabla} \boldsymbol{M}_h^n\|_G^2 - \|\boldsymbol{\nabla} \boldsymbol{M}_h^{n-1}\|_G^2) \end{aligned}$$

for $n > k$, where $\boldsymbol{M}_h^n = [\boldsymbol{m}_h^{n-k+1}, \ldots, \boldsymbol{m}_h^n]$. Next, we aim for an upper bound for the right-hand

side of (3.8.6). As further preparation, we provide bounds for $\boldsymbol{p}_h^n$. Since under the regularity assumptions (1.6.9) we have that $\widehat{\boldsymbol{m}}_h^n \in \boldsymbol{\mathcal{W}}^{1,\infty}(\Omega)$ for $n \geq k$, cf. [7, p. 1022 and (8.5)], we obtain using Lemma 55

$$\|\boldsymbol{p}_h^n\| \leq c\|\widehat{\boldsymbol{m}}_h^n - \widehat{\boldsymbol{m}}_h^{n-1}\|_{\boldsymbol{\mathcal{L}}^\infty}\|\boldsymbol{v}_h^n\| = \mu_n\|\boldsymbol{v}_h^n\| \quad \text{and} \quad \|\boldsymbol{\nabla}\boldsymbol{p}_h^n\| \leq \mu_n\Lambda(\boldsymbol{p}_h^n)\|\boldsymbol{v}_h^n\|$$

for $n > k$, where we introduced the abbreviations

$$\mu_n \coloneqq c\|\widehat{\boldsymbol{m}}_h^n - \widehat{\boldsymbol{m}}_h^{n-1}\|_{\boldsymbol{\mathcal{L}}^\infty}, \qquad \Lambda(\boldsymbol{p}_h^n) \coloneqq \frac{\|\boldsymbol{\nabla}\boldsymbol{p}_h^n\|}{\|\boldsymbol{p}_h^n\|}$$

for $n > k$. Further, we define

$$\nu_n \coloneqq \min\Big\{\frac{\|\boldsymbol{v}_h^n - \boldsymbol{v}_h^{n-1}\|}{\|\boldsymbol{v}_h^n\|}, 1\Big\}$$

for $n > k$. Thus, we obtain the following estimates for the three terms on the right-hand side in (3.8.6) for $n > k$: For the first term

$$\eta\alpha\langle\boldsymbol{v}_h^{n-1}, \boldsymbol{p}_h^n\rangle \leq \mu_n\eta\alpha\|\boldsymbol{v}_h^{n-1}\|\|\boldsymbol{v}_h^n\| \leq \frac{\mu_n\eta\alpha}{2}(\|\boldsymbol{v}_h^{n-1}\|^2 + \|\boldsymbol{v}_h^n\|^2).$$

For the second term, we estimate

$$\begin{aligned}\eta\langle\widehat{\boldsymbol{m}}_h^{n-1} \times \boldsymbol{v}_h^{n-1}, \boldsymbol{v}_h^n - \sigma\boldsymbol{v}_h^{n-1} - \boldsymbol{p}_h^n\rangle &\leq \eta\|\boldsymbol{v}_h^{n-1}\|\Big(\|\boldsymbol{v}_h^n - \sigma\boldsymbol{v}_h^{n-1}\| + \mu_n\|\boldsymbol{v}_h^n\|\Big)\\ &\leq \Big(\frac{\|\boldsymbol{v}_h^n - \sigma\boldsymbol{v}_h^{n-1}\|}{\|\boldsymbol{v}_h^n\|} + \mu_n\Big)\frac{\eta}{2}\Big(\|\boldsymbol{v}_h^{n-1}\|^2 + \|\boldsymbol{v}_h^n\|^2\Big) = (\nu_n + \mu_n)\frac{\eta}{2}\Big(\|\boldsymbol{v}_h^{n-1}\|^2 + \|\boldsymbol{v}_h^n\|^2\Big).\end{aligned}$$

For the third term, we obtain

$$\eta\langle\boldsymbol{\nabla}\boldsymbol{m}_h^{n-1}, \boldsymbol{\nabla}\boldsymbol{p}_h^n\rangle \leq \eta\|\boldsymbol{\nabla}\boldsymbol{m}_h^{n-1}\|\|\boldsymbol{\nabla}\boldsymbol{p}_h^n\| \leq \eta\mu_n\Lambda(\boldsymbol{p}_h^n)\|\boldsymbol{\nabla}\boldsymbol{m}_h^{n-1}\|\|\boldsymbol{v}_h^n\|.$$

If we now also take $\boldsymbol{f}_h^n \neq \boldsymbol{0}$ into account, the right-hand side has the additional term

$$(1 + \mu_n)(\|\boldsymbol{f}_h^n\| + \|\boldsymbol{f}_h^{n-1}\|)\|\boldsymbol{v}_h^n\|.$$

Collecting the above formulas and defining $\eta_{n,\alpha} \coloneqq \frac{1}{2}(1 + \mu_n)\eta\alpha + \frac{1}{2}(\nu_n + \mu_n)\eta$ for $n > k$, we obtain

$$\begin{aligned}&(\alpha - \eta_{n,\alpha})\|\boldsymbol{v}_h^n\|^2 - \eta_{n,\alpha}\|\boldsymbol{v}_h^{n-1}\|^2 + \frac{\ell_{\mathrm{ex}}^2}{\tau_n}(\|\boldsymbol{\nabla}\boldsymbol{M}_h^n\|_{\boldsymbol{G},\Omega}^2 - \|\boldsymbol{\nabla}\boldsymbol{M}_h^{n-1}\|_{\boldsymbol{G},\Omega}^2)\\ &\qquad \leq \ell_{\mathrm{ex}}^2\eta\mu_n\Lambda(\boldsymbol{p}_h^n)\|\boldsymbol{\nabla}\boldsymbol{m}_h^{n-1}\|\|\boldsymbol{v}_h^n\| + (1 + \mu_n)(\|\boldsymbol{f}_h^n\| + \|\boldsymbol{f}_h^{n-1}\|)\|\boldsymbol{v}_h^n\|\end{aligned}$$

for $n > k$. By [7, p. 1028], we have that $\mu_n = O(h)$ using $\tau_n = O(h)$. Clearly, $\nu_n = O(1)$ at least. Finally, $\mu_n\Lambda(\boldsymbol{p}_h^n) = O(1)$, since by inverse estimate $\Lambda(\boldsymbol{p}_h^n) = O(1/h)$. Thus, we may fix h_0 such that $\mu_n\Lambda(\boldsymbol{p}_h^n)$ has a definite (but not necessarily small) bound for $h \leq h_0$. Next,

multiplying by τ_n and summing over $j = k+1$ to n yields

$$\begin{aligned}
&\tau_n(\alpha - \eta_{n,\alpha})\|\boldsymbol{v}_h^n\|^2 + \sum_{j=k+2}^{n} ((\alpha - \eta_{j-1,\alpha})\tau_{j-1} - \eta_{j,\alpha}\tau_j)\|\boldsymbol{v}_h^{j-1}\|^2 + \ell_{\mathrm{ex}}^2\|\boldsymbol{\nabla} \boldsymbol{M}_h^n\|_G^2 \\
&\quad \leq \sum_{j=k+1}^{n} \left\{\tau_j \ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j)\|\boldsymbol{\nabla} \boldsymbol{m}_h^{j-1}\|\|\boldsymbol{v}_h^j\| + \tau_j(1+\mu_j)(\|\boldsymbol{f}_h^j\| + \|\boldsymbol{f}_h^{j-1}\|)\|\boldsymbol{v}_h^j\|\right\} \\
&\qquad + \eta_{k+1,\alpha}\tau_{k+1}\|\boldsymbol{v}_h^k\|^2 + \ell_{\mathrm{ex}}^2\|\boldsymbol{\nabla} \boldsymbol{M}_h^k\|_G^2
\end{aligned}$$

for $n > k$. We first assume $\tau_{j-1} \leq \tau_j \leq \kappa\tau_{j-1}$ for $1 < \kappa < \kappa_{0,k}$ and $j \geq k+2$ to obtain

$$\begin{aligned}
(\alpha - \eta_{j-1,\alpha})\tau_{j-1} - \eta_{j,\alpha}\tau_j &\geq \tau_j\big(\frac{1}{\kappa}(\alpha - \eta_{j-1,\alpha}) - \eta_{j,\alpha}\big) \\
&\geq \tau_j\big(\frac{1}{\kappa_{0,k}}(\alpha - \eta_{j-1,\alpha}) - \eta_{j,\alpha}\big) \\
&\geq \tau_{j-1}\big(\frac{1}{\kappa_{0,k}}(\alpha - \eta_{j-1,\alpha}) - \eta_{j,\alpha}\big),
\end{aligned}$$

if we ensure that $\alpha - \eta_{j-1,\alpha} > 0$ as well as $\omega_j := \frac{1}{\kappa_{0,k}}(\alpha - \eta_{j-1,\alpha}) - \eta_{j,\alpha} > 0$ (for such j), which is the case for α large enough and h small enough. In case $\tau_j < \tau_{j-1}$ we have

$$(\alpha - \eta_{j-1,\alpha})\tau_{j-1} - \eta_{j,\alpha}\tau_j \geq \tau_{j-1}\big(\alpha - \eta_{j-1,\alpha} - \eta_{j,\alpha}\big)$$

for $j \geq k+2$, where $\omega_j := \alpha - \eta_{j-1,\alpha} - \eta_{j,\alpha} > 0$ (for such j) is required. Let $\omega_* > 0$ be the smallest of such weights. Then

$$\begin{aligned}
&\omega_* \sum_{j=k+1}^{n} \tau_j\|\boldsymbol{v}_h^j\|^2 + \ell_{\mathrm{ex}}^2\|\boldsymbol{\nabla} \boldsymbol{M}_h^n\|_G^2 \\
&\quad \leq \sum_{j=k+1}^{n} \left\{\tau_j \ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j)\|\boldsymbol{\nabla} \boldsymbol{m}_h^{j-1}\|\|\boldsymbol{v}_h^j\| + \tau_j(1+\mu_j)(\|\boldsymbol{f}_h^j\| + \|\boldsymbol{f}_h^{j-1}\|)\|\boldsymbol{v}_h^j\|\right\} \\
&\qquad + \ell_{\mathrm{ex}}^2\|\boldsymbol{\nabla} \boldsymbol{M}_h^k\|_G^2 + \eta_{k+1,\alpha}\tau_{k+1}\|\boldsymbol{v}_h^k\|^2
\end{aligned}$$

for $n > k$. On the right-hand side, we proceed with Young's inequality

$$\begin{aligned}
&\sum_{j=k+1}^{n} \left\{\tau_j \ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j)\|\boldsymbol{\nabla} \boldsymbol{m}_h^{j-1}\|\|\boldsymbol{v}_h^j\| + \tau_j(1+\mu_j)(\|\boldsymbol{f}_h^j\| + \|\boldsymbol{f}_h^{j-1}\|)\|\boldsymbol{v}_h^j\|\right\} \\
&\quad \leq \frac{1}{\omega_*} \sum_{j=k+1}^{n} \tau_j(\ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j))^2\|\boldsymbol{\nabla} \boldsymbol{m}_h^{j-1}\|^2 + \frac{2}{\omega_*} \sum_{j=k+1}^{n} \tau_j(1+\mu_j)^2(\|\boldsymbol{f}_h^j\|^2 + \|\boldsymbol{f}_h^{j-1}\|^2) \\
&\qquad + \frac{\omega_*}{2} \sum_{j=k+1}^{n} \tau_j\|\boldsymbol{v}_h^j\|^2
\end{aligned}$$

for $n > k$. Let $\gamma_- > 0$ be the smallest eigenvalue and $\gamma_+ > 0$ be the largest eigenvalue of the

symmetric positive definite matrix G. Then, using the properties of G yields

$$\gamma_-\|\boldsymbol{\nabla} \boldsymbol{m}_h^n\|^2 \leq \|\boldsymbol{\nabla} \boldsymbol{M}_h^n\|_G^2, \qquad \|\boldsymbol{\nabla} \boldsymbol{M}_h^k\|_G^2 \leq \gamma_+ \sum_{i=1}^{k} \|\boldsymbol{\nabla} \boldsymbol{m}_h^i\|^2$$

for $n > k$, cf. [7, p. 1004]. Thus, we obtain

$$\begin{aligned}
&\frac{\omega_*}{2} \sum_{j=k+1}^{n} \tau_j \|\boldsymbol{v}_h^j\|^2 + \ell_{\mathrm{ex}}^2 \gamma_- \|\boldsymbol{\nabla} \boldsymbol{m}_h^n\|^2 \\
&\quad \leq \frac{1}{\omega_*} \sum_{j=k+1}^{n} \tau_j (\ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j))^2 \|\boldsymbol{\nabla} \boldsymbol{m}_h^{j-1}\|^2 + \frac{2}{\omega_*} \sum_{j=k+1}^{n} \tau_j (1+\mu_j)^2 (\|\boldsymbol{f}_h^j\|^2 + \|\boldsymbol{f}_h^{j-1}\|^2) \\
&\qquad + \ell_{\mathrm{ex}}^2 \gamma_+ \sum_{i=1}^{k} \|\boldsymbol{\nabla} \boldsymbol{m}_h^i\|^2 + \eta_{k+1,\alpha} \tau_{k+1} \|\boldsymbol{v}_h^k\|^2
\end{aligned}$$

for $n > k$. We now address the problem in the form

$$\ell_{\mathrm{ex}}^2 \gamma_- \|\boldsymbol{\nabla} \boldsymbol{m}_h^n\|^2 \leq A + \frac{1}{\omega_*} \sum_{j=k+1}^{n} \tau_j (\ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j))^2 \|\boldsymbol{\nabla} \boldsymbol{m}_h^{j-1}\|^2$$

for $n > k$. If we let $\xi = \max_{j=k+1,\dots,n} \{(\ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j))^2 / (\ell_{\mathrm{ex}}^2 \gamma_- \omega_*)\}$, then we obtain by a discrete Gronwall's inequality

$$\begin{aligned}
&\ell_{\mathrm{ex}}^2 \gamma_- \|\boldsymbol{\nabla} \boldsymbol{m}_h^n\|^2 \leq A \mathrm{e}^{\xi T} \\
&\quad = \Big(\frac{2}{\omega_*} \sum_{j=k+1}^{n} \tau_j (1+\mu_j)^2 (\|\boldsymbol{f}_h^j\|^2 + \|\boldsymbol{f}_h^{j-1}\|^2) + \ell_{\mathrm{ex}}^2 \gamma_+ \sum_{i=1}^{k} \|\boldsymbol{\nabla} \boldsymbol{m}_h^i\|^2 + \eta_{k+1,\alpha} \tau_{k+1} \|\boldsymbol{v}_h^k\|^2 \Big) \mathrm{e}^{\xi T}.
\end{aligned}$$

This provides us with an energy bound. We can further derive

$$\begin{aligned}
\frac{\omega_*}{2} \sum_{j=k+1}^{n} \tau_j \|\boldsymbol{v}_h^j\|^2 &\leq A + \frac{1}{\omega_*} \sum_{j=k+1}^{n} \tau_j (\ell_{\mathrm{ex}}^2 \eta \mu_j \Lambda(\boldsymbol{p}_h^j))^2 \|\boldsymbol{\nabla} \boldsymbol{m}_h^{j-1}\|^2 \\
&\leq A(1 + \xi T \mathrm{e}^{\xi T}) \leq A(1+\xi T) \mathrm{e}^{\xi T}.
\end{aligned}$$

Lastly, testing (3.8.1) for $n = k$ with $\boldsymbol{\varphi}_h^k = \boldsymbol{v}_h^k$, i.e., $\alpha \langle \boldsymbol{v}_h^k, \boldsymbol{v}_h^k \rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^k, \boldsymbol{\nabla} \boldsymbol{v}_h^k \rangle = \langle \boldsymbol{f}_h^k, \boldsymbol{v}_h^k \rangle$, and applying the definition of the BDF(k) scheme (3.8.3) together with Young's inequalities, we obtain

$$\begin{aligned}
\tau_{k+1} \frac{1}{2\kappa_{k+1}} \alpha \|\boldsymbol{v}_h^k\|^2 + \ell_{\mathrm{ex}}^2 \frac{\delta_0}{2} \|\boldsymbol{\nabla} \boldsymbol{m}_h^k\|^2 &\leq \frac{\ell_{\mathrm{ex}}^2}{2\delta_0} \Big(\sum_{j=0}^{k-1} |\delta_{k-j}| \|\boldsymbol{\nabla} \boldsymbol{m}_h^j\| \Big)^2 + \frac{\tau_k}{2\alpha} \|\boldsymbol{f}_h^k\|^2 \\
&\leq C \sum_{j=0}^{k-1} \|\boldsymbol{\nabla} \boldsymbol{m}_h^j\|^2 + \frac{\tau_k}{2\alpha} \|\boldsymbol{f}_h^k\|^2,
\end{aligned}$$

which concludes the proof. □

Next, we use the DOC kernels introduced in Section 3.7 to derive an $\mathcal{L}^2$-bound for the

solution of (3.8.1) from the discrete energy inequality of Theorem 56.

Theorem 57 ($\mathcal{L}^2$-bound for orders $k = 1, 2, 3$)**.** *Consider the discretization* (3.8.1) *of the LLG equation* (1.2.16) *for* $k \in \{1, 2, 3\}$ *with finite elements of polynomial degree* $p \geq 1$*. We assume* $\tau = \mathcal{O}(h)$ *and, for* $k \geq 2$*, that the time-steps ratios satisfy* $\kappa_i = \tau_i/\tau_{i-1} \leq \kappa_0 < \kappa_{0,k}$ *for all* $2 \leq i \leq n$*, where*

$$\kappa_{0,2} = 1 + \sqrt{2}, \qquad \kappa_{0,3} \approx 1.476,$$

and $\alpha \geq \alpha_{0,k} > 0$ *for some* $\alpha_{0,k}$ *that depends on* $\kappa_{0,k}$*. Then, the numerical solution satisfies the following* $\mathcal{L}^2$*-bound: For* $n > k$ *it holds*

$$\|\boldsymbol{m}_h^n\|^2 \leq C \left(\sum_{i=0}^{k-1} \|\boldsymbol{m}_h^i\|_{\boldsymbol{\mathcal{H}}^1}^2 + \sum_{j=k}^{n} \tau_j \|\boldsymbol{f}_h^j\|^2 \right)$$

under the regularity assumptions (1.6.9)*, where the constant* $C > 0$ *is independent of the spatial mesh size* h *and the time-steps* τ_j*, but depends exponentially on* T*.*

Proof. The idea of the proof is to connect $\|\boldsymbol{m}_h^n\|^2$ to $\sum_{j=k}^n \tau_j \|\boldsymbol{v}_h^j\|^2$, while utilizing the statement of Theorem 56. For the convenience of the reader, we recall (3.7.6)

$$\boldsymbol{m}_h^n - \boldsymbol{m}_h^{k-1} + \sum_{\ell=k}^{n} \mathcal{I}_k^\ell = \sum_{\ell=k}^{n} \left(\tau_\ell \boldsymbol{v}_h^\ell \sum_{j=\ell}^{n} \theta_{j-\ell}^{(k,j)} \right)$$

for $n \geq k$, where

$$\mathcal{I}_k^\ell = \sum_{j=1}^{k-1} \nabla_\tau \boldsymbol{m}_h^j \sum_{i=k}^{\ell} \theta_{\ell-i}^{(k,\ell)} b_{i-j}^{(k,i)}$$

for $\ell \geq k$.

$\boldsymbol{k = 1}$**:** We immediately obtain, for $n \geq 1$,

$$\|\boldsymbol{m}_h^n\|^2 \leq \|\boldsymbol{m}_h^0\|^2 + \sum_{\ell=1}^{n} \tau_\ell^2 \|\boldsymbol{v}_h^\ell\|^2.$$

$\boldsymbol{k = 2}$**:** Using the explicit formula of the DOC kernels (3.7.8), we obtain

$$\sum_{j=\ell}^{n} \theta_{j-\ell}^{(2,j)} = \sum_{j=\ell}^{n} \frac{1+\kappa_\ell}{1+2\kappa_\ell} \prod_{i=\ell+1}^{j} \frac{\kappa_i^2}{1+2\kappa_i}$$

for $n \geq \ell \geq 2$. We have $\frac{1+\kappa_i}{1+2\kappa_i} \leq 1$ for $\kappa_i < \kappa_{0,2} = 1 + \sqrt{2}$ and $2 \leq i \leq n$. Further, let $\kappa_\star$ be the maximal value of all step ratios κ_i, i.e., $\kappa_i \leq \kappa_\star < \kappa_{0,2} = 1 + \sqrt{2}$ for all $2 \leq i \leq n$. By computing the partial geometric series, we obtain

$$\sum_{j=\ell}^{n} \theta_{j-\ell}^{(2,j)} = \sum_{j=\ell}^{n} \frac{1+\kappa_\ell}{1+2\kappa_\ell} \prod_{i=\ell+1}^{j} \frac{\kappa_i^2}{1+2\kappa_i} \leq \sum_{j=\ell}^{n} \left(\frac{\kappa_\star^2}{1+2\kappa_\star} \right)^{k-j} \leq \frac{1+2\kappa_\star}{1+2\kappa_\star-\kappa_\star^2} =: C_{\kappa_\star} \tag{3.8.7}$$

for $n \geq \ell \geq 2$. Using the above inequalities, we end up with

$$\|\boldsymbol{m}_h^n\|^2 \leq \|\boldsymbol{m}_h^1\|^2 + \sum_{\ell=2}^{n} C_{\kappa_\star}^2 \tau_\ell^2 \|\boldsymbol{v}_h^\ell\|^2 + \|\sum_{\ell=2}^{n} \mathcal{I}_k^\ell\|^2$$

for $n \geq 2$. By the definition of the DC kernels (3.7.3), we have that $|b_1^{(2,2)}| = \kappa_2^2/(1+\kappa_2) \leq 2$ for $\kappa_2 \leq 1+\sqrt{2}$ and $b_j^{(2,n)} = 0$ for $j \geq 2$. Thus, applying (3.8.7) yields

$$\|\sum_{\ell=2}^{n} \mathcal{I}_2^\ell\|^2 = \|\nabla_\tau \boldsymbol{m}_h^1 \sum_{j=2}^{n} \theta_{j-2}^{(2,j)} b_1^{(2,2)}\|^2 \leq 4 C_{\kappa_\star}^2 \|\nabla_\tau \boldsymbol{m}_h^1\|^2.$$

***k* = 3:** Utilizing the bound of the DOC kernels from Lemma 52 leads to

$$\|\boldsymbol{m}_h^n\|^2 \leq \|\boldsymbol{m}_h^2\|^2 + \sum_{\ell=3}^{n} C_3^2 \tau_\ell^2 \|\boldsymbol{v}_h^\ell\|^2 + \|\sum_{\ell=3}^{n} \mathcal{I}_3^\ell\|^2$$

for $n \geq 3$. Since

$$\mathcal{I}_3^\ell = \nabla_\tau \boldsymbol{m}_h^1 \theta_{\ell-3}^{(3,\ell)} b_2^{(3,3)} + \nabla_\tau \boldsymbol{m}_h^2 \Big(\theta_{\ell-3}^{(3,\ell)} b_1^{(3,3)} + \theta_{\ell-4}^{(3,\ell)} b_2^{(3,4)} \Big),$$

we obtain by using again the bound of the DOC kernels from Lemma 52

$$\|\sum_{\ell=3}^{n} \mathcal{I}_3^\ell\|^2 \leq C_3^2 \Big(|b_2^{(3,3)}|^2 \, \|\nabla_\tau \boldsymbol{m}_h^1\|^2 + (|b_1^{(3,3)}|^2 + |b_2^{(3,4)}|^2) \, \|\nabla_\tau \boldsymbol{m}_h^2\|^2 \Big).$$

The proof is completed by inserting the above inequalities for $k = 1, 2, 3$ into the discrete energy inequality. □

Remark 58. *In the proof of Lemma 57, we proved, under the assumption $\kappa_i < \kappa_{0,2} = 1+\sqrt{2}$, the upper bound for the DOC kernels*

$$\sum_{j=\ell}^{n} \theta_{j-\ell}^{(2,j)} < C_{\kappa_\star}$$

using the explicit formula of the DOC kernels (3.7.8) *for the* BDF(2) *scheme. This result can be generalized to cover finitely many step ratios that do not satisfy the upper bound $1+\sqrt{2}$. More precisely, it is possible to choose $\kappa_i \in [1+\sqrt{2}, \frac{3+\sqrt{17}}{2}]$ if the size of the index set $\overline{K} = \{i \,:\, \kappa_i \in [1+\sqrt{2}, \frac{3+\sqrt{17}}{2}]\}$ is relatively small in the sense that $|\overline{K}| \ll n$, cf. [93, Lemma 3.3]. However, this generalization is primarily of theoretical interest. For practical applications, the recommended step-size relation is still the classical result from [77] given here, see [93, p. 1223].*

CHAPTER 4

Space and time adaptive algorithm

In this chapter we establish a variable step-size variable order (VSVO) time-stepping scheme and integrate it into a space and time adaptive framework for the LLG equation. We begin with a rigorous derivation of the local truncation error (LTE) for the backward differentiation formulas. This derivation leads to the development of computable LTE approximations that serve as the foundation for the adaptive time-step size control. By dynamically adjusting the time-step size, we ensure that the temporal error remains below a prescribed tolerance throughout the entire simulation. We further describe an order-selection strategy that is designed to prioritize numerical stability while determining the most efficient degree of the BDF schemes. We conclude the chapter with a comprehensive description of the space and time adaptive algorithm. We demonstrate the performance and robustness of the solver through numerical experiments in Chapter 5.

4.1 Estimating the local truncation error

In this section, we consider the partition

$$[0 = t_0, t_1, t_2, \dots, t_{k+1}] = [0, \tau_1, \tau_{1,2}, \dots, \tau_{1,\dots,k+1}] \tag{4.1.1}$$

for the variable step-size BDF(k) method with $k \in \mathbb{N}$, where the abbreviations $\tau_{1,\dots,k}$ are defined in (3.1.10). Compared with the time mesh in (3.1.8), this partition contains one additional node, as it is required later for approximating higher order derivatives. By an index shift, the partition can be translated into $[t_{n-k}, \dots, t_{n+1}]$ for $n \geq k$. Furthermore, let $\tau = \max\{\tau_1, \dots, \tau_{k+1}\}$ denote the maximal time-step size. We consider the ordinary differential equation $u'(t) = f(t, u(t))$ for $t \in [0, t_{k+1}]$ as in (3.1.1) with the notation $u^i = u(t_i)$ for $i = 0, \dots, k+1$.

Next, we derive the local truncation error for each BDF(k) scheme. To this end, we expand each u^{k-i} for $i = 1, \dots, k$ in a Taylor series about t_k (cf. (3.1.12)) and insert these into Φ^k defined in (3.1.11). For $k = 1$ we obtain

$$\begin{aligned} \Phi^1(u, t_2) &= u^2 - u^1 - \tau_2 u'(t_2) \\ &= u^2 - (u^2 - \tau_2 u'(t_2) + \frac{\tau_2^2}{2} u''(t_2)) - \tau_2 u'(t_2) + \mathcal{O}(\tau^3) \\ &= -\frac{\tau_2^2}{2} u''(t_2) + \mathcal{O}(\tau^3). \end{aligned} \tag{4.1.2}$$

For $k = 2$, with $\kappa = \kappa_3 = \tau_3/\tau_2$ and the coefficients of the variable step-size BDF(2) method

given in (3.4.2), we derive

$$\begin{aligned}\Phi^2(u,t_3) &= \frac{1+2\kappa}{1+\kappa}u^3 - (1+\kappa)u^2 + \frac{\kappa^2}{1+\kappa}u^1 - \tau_3 u'(t_3)\\ &= (1+\kappa)\frac{\tau_3^3}{6}u'''(t_3) - \frac{\kappa^2}{1+\kappa}\frac{\tau_{2,3}^3}{6}u'''(t_3) + \mathcal{O}(\tau^4)\\ &= -\frac{\tau_3^2\tau_{2,3}}{6}u'''(t_3) + \mathcal{O}(\tau^4).\end{aligned}$$

Similarly, we find for $k=3$

$$\begin{aligned}\Phi^3(u,t_4) &= \frac{1}{24}\Big(\tau_4^4\delta_1 + \tau_{3,4}^4\delta_2 + \tau_{2,3,4}^4\delta_3\Big)u''''(t_4) + \mathcal{O}(\tau^5)\\ &= -\frac{\tau_4^2\tau_{3,4}\tau_{2,3,4}}{24}u''''(t_4) + \mathcal{O}(\tau^5),\end{aligned}$$

where the coefficients δ_1, δ_2 and δ_3 of the variable step-size BDF(3) method are given in (3.1.14). In summary, the leading term of the local truncation error T_k for $k=1,2,3$ is given by

$$T_1 \coloneqq \frac{\tau_2^2}{2}u''(t_2), \qquad T_2 \coloneqq \frac{\tau_3^2\tau_{2,3}}{6}u'''(t_3), \qquad T_3 \coloneqq \frac{\tau_4^2\tau_{3,4}\tau_{2,3,4}}{24}u''''(t_4). \tag{4.1.3}$$

For an equidistant partition, the leading term of the local truncation error simplifies to

$$T_k = \frac{1}{k+1}\tau^{k+1}u^{(k+1)}(t_{k+1}). \tag{4.1.4}$$

4.1.1 Approximation with finite-differences

To evaluate the leading term of the local truncation error a posteriori, we need approximations for the higher-order time derivatives. A natural approach is to obtain these derivatives by finite-differences constructed using the Lagrange interpolation polynomial.

Thus, let $P = P[u|t_0,\dots,t_{k+1}] \in \mathbb{P}_{k+1}$ denote the Lagrange interpolation polynomial defined in (3.1.15). Then, with the derivative of order $k+1$ of P, we define the consistent finite-difference approximation for the $(k+1)$-th time derivative by

$$\overline{\partial}_\tau^{k+1}u^{k+1} \coloneqq \partial_t^{k+1}P[u|t_0,\dots,t_{k+1}], \tag{4.1.5}$$

which is constant in t but depends on the time-steps $\tau_1,\dots\tau_{k+1}$. For $k=1$, this yields the explicit formula

$$\overline{\partial}_\tau^2 u^2 \coloneqq \partial_t^2 P[u|t_0,t_1,t_2] = 2\Big(\frac{1}{\tau_{1,2}\tau_2}u^2 - \frac{1}{\tau_1\tau_2}u^1 + \frac{1}{\tau_1\tau_{1,2}}u^0\Big) = \frac{2\kappa_2}{1+\kappa_2}\overline{\partial}^2 u^2, \tag{4.1.6}$$

where $\overline{\partial}^2$ is defined in (2.6.1) for variable time-steps. Further, for $k = 2$, we have

$$\begin{aligned}\overline{\partial}_\tau^3 u^3 &:= \partial_t^3 P[u|t_0, t_1, t_2, t_3] = \frac{3}{\tau_{1,2,3}}\Big(\overline{\partial}_\tau^2 u^3 - \overline{\partial}_\tau^2 u^2\Big) \\ &= 6\Big(\frac{1}{\tau_{1,2,3}\tau_{2,3}\tau_3}u^3 - \frac{1}{\tau_{1,2}\tau_2\tau_3}u^2 + \frac{1}{\tau_1\tau_2\tau_{2,3}}u^1 - \frac{1}{\tau_1\tau_{1,2}\tau_{1,2,3}}u^0\Big).\end{aligned} \tag{4.1.7}$$

Similarly, for $k = 3$, we obtain

$$\begin{aligned}\overline{\partial}_\tau^4 u^4 := \partial_t^4 P[u|t_0, \dots, t_4] = 24\Big(&\frac{1}{\tau_{1,2,3,4}\tau_{2,3,4}\tau_{3,4}\tau_4}u^4 - \frac{1}{\tau_{1,2,3}\tau_{2,3}\tau_3\tau_4}u^3 + \frac{1}{\tau_{1,2}\tau_2\tau_3\tau_{3,4}}u^2 \\ &- \frac{1}{\tau_1\tau_2\tau_{2,3}\tau_{2,3,4}}u^1 + \frac{1}{\tau_1\tau_{1,2}\tau_{1,2,3}\tau_{1,2,3,4}}u^0\Big).\end{aligned}$$

Thus, using (4.1.5) we can approximate the local truncation error given in (4.1.3) by

$$T_1 \approx \frac{\tau_2^2}{2}\overline{\partial}_\tau^2 u(t_2), \qquad T_2 \approx \frac{\tau_3^2\tau_{2,3}}{6}\overline{\partial}_\tau^3 u(t_3), \qquad T_3 \approx \frac{\tau_4^2\tau_{3,4}\tau_{2,3,4}}{24}\overline{\partial}_\tau^4 u(t_4). \tag{4.1.8}$$

In the context of the LLG equation, the first-order derivative is already computed via the tangent plane scheme, cf. Section 1.6. Thus, we want to exploit this by using finite-difference approximations of one order lower in (4.1.8). Let $v^i = v(t_i) := u'(t_i)$ for $i = 0, \dots, k+1$. Assume that the approximations u^i and v^i are available for $i = 0, \dots, k$ and that we aim to compute the approximation u^{k+1} with the step-size τ_{k+1}. Having an approximation of the leading term of the local truncation error T_k by (4.1.8) at hand, we want to determine the next time-step size for a given tolerance TOL such that

$$\mathrm{TOL} = |T_k|.$$

Thus, for $k = 1$, we have

$$\tau_2 = \left(\frac{2\,\mathrm{TOL}}{|\overline{\partial}_\tau^2 u(t_2)|}\right)^{1/2}.$$

Similarly, for $k = 2, 3$, we obtain

$$\tau_3 = \left(\frac{6\,\mathrm{TOL}}{\tau_{2,3}|\overline{\partial}_\tau^3 u(t_3)|}\right)^{1/2}, \qquad \tau_4 = \left(\frac{24\,\mathrm{TOL}}{\tau_{3,4}\tau_{2,3,4}|\overline{\partial}_\tau^4 u(t_4)|}\right)^{1/2}.$$

Since u^{k+1} is unknown while choosing τ_{k+1}, we estimate the required $(k+1)$-st derivative of u by the k-th derivative of $v = u'$ by

$$\overline{\partial}_\tau^{k+1} u(t_{k+1}) \approx \overline{\partial}_\tau^k v(t_{k+1}).$$

Next, by Taylor expanding $\overline{\partial}_\tau^k v(t_{k+1})$ about t_k, we obtain

$$\overline{\partial}_\tau^k v(t_{k+1}) = \overline{\partial}_\tau^k v(t_k) + \mathcal{O}(\tau) \approx \overline{\partial}_\tau^k v(t_k).$$

In summary, we compute the next time-step size τ_{k+1} by

$$\tau_2 = \left(\frac{2\,\mathrm{TOL}}{|\overline{\partial}_\tau^1 v(t_1)|}\right)^{1/2}, \qquad \tau_3 = \left(\frac{6\,\mathrm{TOL}}{\tau_{2,3}|\overline{\partial}_\tau^2 v(t_2)|}\right)^{1/2}, \qquad \tau_4 = \left(\frac{24\,\mathrm{TOL}}{\tau_{3,4}\tau_{2,3,4}|\overline{\partial}_\tau^3 v(t_3)|}\right)^{1/2}. \tag{4.1.9}$$

Alternatively, we can proceed as in (3.1.16) and interpolate the recent values u^i and v^i for $i = 0, \dots, k$ onto a locally uniform mesh and compute the finite-differences on the uniform partition.

4.1.2 Approximation via extrapolation

As an alternative to the finite-difference approach, we can follow the idea of [43, 113]. For this part, we assume that we have an equidistant partition. This partition is obtained by either performing k consecutive time-steps with the constant step-size τ or by interpolating the recent values onto a locally uniform time mesh. Under this assumption, the equidistant formula (4.1.4) can be used to approximate the leading term of the local truncation error by

$$T_k = \frac{1}{k+1}\tau^{k+1}u^{(k+1)}(t_{k+1}) \approx \frac{1}{k+1}\tau^{k+1}\overline{\partial}_\tau^{k+1}u^{k+1} = \frac{1}{k+1}\nabla_\tau^{k+1}u^{k+1}, \tag{4.1.10}$$

where the discrete difference ∇_τ^{k+1} is defined in (3.1.4). Resolving each discrete difference for u^{k+1} leads to

$$\frac{1}{k+1}\nabla_\tau^{k+1}u^{k+1} = \frac{1}{k+1}\big(u^{k+1} - \sum_{i=0}^{k}\nabla_\tau^i u^k\big). \tag{4.1.11}$$

Recall the interpolation polynomial of $u^0, \dots, u^k$ from (3.1.5), namely

$$p(t) = \sum_{j=0}^{k}(-1)^j\binom{-s}{j}\nabla_\tau^j u^k \qquad \text{with} \qquad s = \frac{t - t_k}{\tau}$$

for $t \in \mathbb{R}$. With the explicit representation of the binomial coefficient

$$\binom{-1}{j} = \frac{(-1)(-2)\cdots(-j)}{j!} = (-1)^j$$

for $j \geq 0$, we can evaluate p at t_{k+1} explicitly by

$$\widehat{u}^{k+1} \coloneqq p(t_{k+1}) = \sum_{i=0}^{k}\nabla_\tau^i u^k. \tag{4.1.12}$$

Finally, inserting (4.1.12) and (4.1.11) into (4.1.10) yields the approximation

$$T_k \approx \frac{1}{k+1}\big(u^{k+1} - \sum_{i=0}^{k}\nabla_\tau^i u^k\big) = \frac{1}{k+1}\big(u^{k+1} - \widehat{u}^{k+1}\big) \tag{4.1.13}$$

for the leading term of the local truncation error.

To determine an appropriate step-size for a prescribed tolerance TOL, we assume that the solution u^n at t_n has been computed using BDF(k) with the step-size τ_n on an equidistant partition. In this case, the local truncation error at t_n is given by

$$\frac{1}{k+1}\tau_n^{k+1}u^{(k+1)}(t_n) + \mathcal{O}(\tau_n^{k+2}) =: \tau_n^{k+1}\phi(t_n) + \mathcal{O}(\tau_n^{k+2}),$$

cf. (4.1.4). We want to compute u^{n+1} with the step-size $\tau_{n+1} = \sigma\tau_n$ such that the local truncation error at t_{n+1} is approximately equal to the prescribed tolerance TOL. The local truncation error at t_{n+1} is given by

$$(\sigma\tau_n)^{k+1}\phi(t_{n+1}) + \mathcal{O}((\sigma\tau_n)^{k+2}).$$

Using the Taylor expansion of $\phi(t_{n+1})$ about t_n

$$\phi(t_{n+1}) = \phi(t_n) + \sigma\tau_n\phi'(t_n) + \mathcal{O}((\sigma\tau_n)^2),$$

we conclude

$$(\sigma\tau_n)^{k+1}\phi(t_{n+1}) + \mathcal{O}((\sigma\tau_n)^{k+2}) = \sigma^{k+1}\tau_n^{k+1}\phi(t_n) + \mathcal{O}((\sigma\tau_n)^{k+2}) \approx \sigma^{k+1}T_k^n,$$

where $T_k^n = \frac{1}{k+1}(u^n - \widehat{u}^n)$ is the leading term of the local truncation error given in (4.1.13) evaluated at t_n. Thus, we determine σ such that the local truncation error at t_{n+1} is approximately equal to the tolerance, namely $\mathrm{TOL} = |\sigma^{k+1}T_k^n|$, by

$$\sigma = \left(\frac{\mathrm{TOL}}{|T_k^n|}\right)^{\frac{1}{k+1}} = \left(\frac{(k+1)\,\mathrm{TOL}}{|u^n - \widehat{u}^n|}\right)^{\frac{1}{k+1}}. \tag{4.1.14}$$

Finally, we obtain the next step-size by $\tau_{n+1} = \sigma\tau_n$.

4.1.3 Time-step selection for residual-based estimators

For the BDF(2) discretization of the LLG equation we have a posteriori control for the $\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{H}}^1(\Omega))$-error according to Theorem 19 for uniform time-step sizes and Theorem 29 for variable time-step sizes. To exploit these bounds in an adaptive time-stepping strategy, we aim to equilibrate the local contribution of the residual-based estimators introduced in Definition 18 and Definition 28, respectively. This approach is well suited for controlling the $\mathcal{L}^\infty$-error in time [88]. Thus, we define the local error indicators

$$\varepsilon_{1,n} \coloneqq \tau_n^4\|\boldsymbol{\nabla}\overline{\partial}_n^2\boldsymbol{m}_h^n\|^2, \tag{4.1.15}$$

$$\varepsilon_{2,n} \coloneqq \frac{\tau_n^2}{\kappa_n^2}\|\frac{\kappa_n}{\kappa_n+1}\overline{\partial}^2\boldsymbol{m}_h^n - \frac{\kappa_{n-1}}{\kappa_{n-1}+1}\overline{\partial}^2\boldsymbol{m}_h^{n-1}\|^2 \tag{4.1.16}$$

for $n \geq 2$, cf. Definition 28. The relation to the global error indicators is then given by

$$\widetilde{\mathcal{E}}_1 = \max_{2\leq n\leq N} \varepsilon_{1,n}, \qquad \widetilde{\mathcal{E}}_2 = \sum_{n=2}^{N} \tau_n \varepsilon_{2,n}.$$

By Theorem 29, the $\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{H}}^1(\Omega))$-error is bounded, among other error indicators, by $\widetilde{\mathcal{E}}_1, \widetilde{\mathcal{E}}_2$. Consequently, a suitable adaptive strategy is to choose the next time-step size with a prescribed tolerance TOL, such that the local contributions are controlled. Similar to (4.1.14) with $k = 2$, we update the time-step size by

$$\tau_{n+1} = \left(\frac{\mathrm{TOL}}{\sqrt{\varepsilon_{i,n}}}\right)^{\frac{1}{3}} \tau_n \tag{4.1.17}$$

for $n \geq 2$ with either $i = 1$ or $i = 2$.

4.1.4 Time-step size restrictions

In general, after computing the next time-step size, for instance by (4.1.9), (4.1.14) or (4.1.17), we need to ensure that the ratio of consecutive time-step sizes is reasonably bounded. This is in particular important for the BDF(k) method with $k \geq 2$, to ensure stability in the sense of the discrete energy inequality (cf. Theorem 56). Thus, each time-step size selection is subject to a restriction with respect to the previous step-size with ratio bounds $\underline{\kappa}_k$ and $\overline{\kappa}_k$ depending on the order k in the form

$$\tau_{n+1} = \max\left\{\tau_{\min},\, \underline{\kappa}_k \tau_n,\, \min\{\tau^\star_{n+1},\, \overline{\kappa}_k \tau_n,\, \tau_{\max}\}\right\}. \tag{4.1.18}$$

Here, $\tau^\star_{n+1}$ is the time-step size computed by (4.1.9), (4.1.14) or (4.1.17), τ_{n+1} is the actual choice and $\tau_{\min} > 0$ and $\tau_{\max} > 0$ are prescribed lower and upper bounds for the time-step size. For instance, we take $\tau_{\min} = 10^{-10}$ and $\tau_{\max} = T/10$, where $T > 0$ is the final simulation time. In general, we employ $\underline{\kappa}_k = \frac{1}{4}$ for $k = 1, 2, 3, 4$ and

$$\overline{\kappa}_1 = 4, \qquad \overline{\kappa}_2 = 2.4, \qquad \overline{\kappa}_3 = 1.6, \qquad \overline{\kappa}_4 = 1.2\,. \tag{4.1.19}$$

In the case that the estimated error is close or equal to zero, the time-step choice is controlled by the maximum cutoff value $\tau_{\max}$ in (4.1.18).

4.2 Computing the predictor

In this section, we discuss the construction of the predictor required for the tangent plane scheme described in Section 1.6. We present a practical choice, applicable to arbitrary time meshes. Note that in the proof of the a posteriori error estimate, Theorem 19, no assumptions on the predictor are imposed, cf. Section 2.3.

Assume that we have $\ell + 1$, $\ell \in \mathbb{N}$, previously computed values $\boldsymbol{m}_h^{n-\ell}, \ldots, \boldsymbol{m}_h^n$ at time

$t_{n-\ell}, \dots, t_n$. We construct the Lagrange interpolation polynomial of degree ℓ by

$$\boldsymbol{P}^\ell(t) \coloneqq \boldsymbol{P}^\ell[\boldsymbol{m}_h^{n-\ell}, \dots, \boldsymbol{m}_h^n](t) \coloneqq \sum_{j=0}^{\ell} \boldsymbol{m}_h^{n-j} L_{n-j}^n(t), \tag{4.2.1}$$

where L_{n-j}^n denotes the Lagrange basis polynomial associated with $t_{n-\ell}, \dots, t_n$, cf. (3.1.15). Equivalently, $\boldsymbol{P}^\ell$ can be expressed in Newton form by

$$\boldsymbol{P}^\ell(t) = \boldsymbol{m}_h[t_n] + \sum_{j=1}^{\ell} \boldsymbol{m}_h[t_n, \dots, t_{n-j}] \prod_{i=0}^{j-1} (t - t_{n-i})$$

with the recursively defined divided differences

$$\boldsymbol{m}_h[t_n, \dots, t_{n-j}] \coloneqq \frac{\boldsymbol{m}_h[t_n, \dots, t_{n-j+1}] - \boldsymbol{m}_h[t_{n-1}, \dots, t_{n-j}]}{t_n - t_{n-j}}$$

for $1 \le j \le n$ and $\boldsymbol{m}_h[t_j] \coloneqq \boldsymbol{m}_h^j$ for $j \ge 0$. For the BDF(k) method with $k \ge 1$ the predictor $\widehat{\boldsymbol{m}}_h^{n+1}$ at t_{n+1} is then computed by

$$\widehat{\boldsymbol{m}}_h^{n+1} \coloneqq \frac{\boldsymbol{P}^k(t_{n+1})}{|\boldsymbol{P}^k(t_{n+1})|}. \tag{4.2.2}$$

This choice, however, requires one additional historical value compared to the BDF(k) method itself. This predictor differs from the choice in [7, (2.1)] for uniform time meshes, where the authors take

$$\widehat{\boldsymbol{m}}_h^{n+1} \coloneqq \frac{\boldsymbol{P}^{k-1}(t_{n+1})}{|\boldsymbol{P}^{k-1}(t_{n+1})|}.$$

In the case of a uniform time mesh, the polynomial $\boldsymbol{P}^\ell$ can be equivalently expressed as

$$\boldsymbol{P}^\ell(t) = \sum_{j=0}^{\ell} \gamma_j^{(\ell)} \boldsymbol{m}_h^{n-j},$$

where the coefficients $\gamma_j^{(\ell)}$ are those of the polynomial

$$\gamma(z) = \frac{1}{z}(1 - (1-z)^{\ell+1}) = \sum_{j=0}^{\ell} \gamma_j^{(\ell)} z^j.$$

In practice, we normalize the predictor (4.2.2) for each spatial degree of freedom. For instance, in the case of linear finite elements, we normalize in all mesh nodes.

4.3 GRADIENT RECOVERY ESTIMATOR

Let $\Omega \subset \mathbb{R}^d$. Denote by $V_h = \mathcal{S}_h^p(\mathcal{T}_h)$ the space of globally continuous piecewise polynomials of degree p on the mesh $\mathcal{T}_h$ and set $\boldsymbol{V}_h = V_h^d$. The main idea is to construct a post-processing operator

$$\boldsymbol{G}_h : \mathcal{H}^1(\Omega) \to \boldsymbol{V}_h,$$

that yields a good approximation of the exact gradient. This approach is commonly known as gradient recovery or post-processed gradient [4, 5, 119]. If we suppose that for a constant $0 \leq \beta < 1$ the saturation assumption

$$\|\boldsymbol{\nabla} u - \boldsymbol{G}_h(u_h)\| \leq \beta \|\boldsymbol{\nabla} u - \boldsymbol{\nabla} u_h\|$$

is satisfied for the operator $\boldsymbol{G}_h$, we immediately obtain the equivalence of the easily computable gradient recovery error to the $\mathcal{H}^1$-error by

$$\frac{1}{1+\beta}\|\boldsymbol{\nabla} u_h - \boldsymbol{G}_h(u_h)\| \leq \|\boldsymbol{\nabla} u - \boldsymbol{\nabla} u_h\| \leq \frac{1}{1-\beta}\|\boldsymbol{\nabla} u_h - \boldsymbol{G}_h(u_h)\|.$$

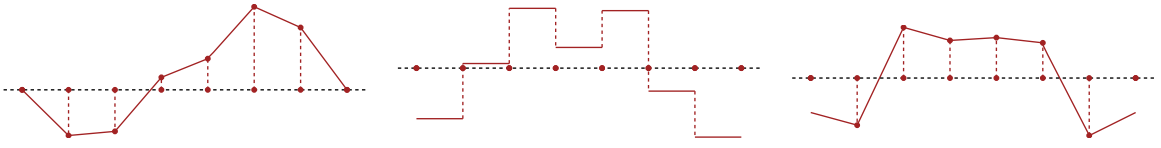
This motivates the choice for the local error indicator

$$\eta_K^{\boldsymbol{G}}(u_h) = \|\boldsymbol{\nabla} u_h - \boldsymbol{G}_h(u_h)\|_{\mathcal{L}^2(K)} \tag{4.3.1}$$

for each cell $K \in \mathcal{T}_h$. For instance, we take $\boldsymbol{G}_h$ as the $\mathcal{L}^2$-projection of the gradient of the finite element function u_h onto $\boldsymbol{V}_h$ by

$$\langle \boldsymbol{G}_h(u_h), \boldsymbol{\varphi}_h \rangle = \langle \boldsymbol{\nabla} u_h, \boldsymbol{\varphi}_h \rangle$$

for all $\boldsymbol{\varphi}_h \in \boldsymbol{V}_h$. This post-processing step is illustrated in Figure 4.3.1. Note that computing the $\mathcal{L}^2$-projection requires solving a linear system with the mass matrix.



(a) $u_h \in \mathcal{S}_h^1(\mathcal{T}_h)$ (b) $\boldsymbol{\nabla} u_h \in \widetilde{\mathcal{S}}_h^0(\mathcal{T}_h)$ (c) $\boldsymbol{G}_h(u_h) \in \mathcal{S}_h^1(\mathcal{T}_h)$

Figure 4.3.1: Gradient recovery for linear finite elements in one dimension. The discrete finite element function u_h is given in (a). Its gradient, belonging to the broken finite element space of piecewise constants denoted by $\widetilde{\mathcal{S}}_h^0(\mathcal{T}_h)$, is illustrated in (b). The recovery operator $\boldsymbol{G}_h$ maps the gradient $\boldsymbol{\nabla} u_h$ into the finite element space V_h, shown in (c).

Let $\mathcal{I}_h^p : C^0(\Omega) \to \boldsymbol{V}_h$ denote the nodal interpolation operator. If the recovery operator $\boldsymbol{G}_h$ satisfies a consistency, localization and boundedness condition, then the recovered gradient of an interpolated function $\mathcal{I}_h^p u$ provides a reliable approximation of the exact gradient $\boldsymbol{\nabla} u$.

More precisely, under the assumptions of [4, Lemma 2.4], $\boldsymbol{G}_h$ satisfies the a priori estimate

$$\|\boldsymbol{\nabla} u - \boldsymbol{G}_h(\mathcal{I}_h^p u)\| \leq C h^{p+1} \, |u|_{\mathcal{H}^{p+2}(\Omega)}$$

for $u \in \mathcal{H}^{p+2}(\Omega)$. If additionally a superconvergence property holds, we obtain that the recovery operator is asymptotically exact [4, Lemma 2.5]. In practice, however, superconvergence occurs only in very special circumstances and is not expected on adaptive meshes. In [119, Thm. 1.31], the equivalence of an averaging gradient recovery operator to the face residuals error estimator is established for linear finite elements.

4.4 Marking for mesh refinement and coarsening

Let $\theta_\mathrm{r}, \theta_\mathrm{c} \in [0, 1)$ be two constants that indicate the percentage of cells that we aim to refine or coarsen by the bulk criterion [59]. Suppose that a local error indicator η_K has been computed for each cell $K \in \mathcal{T}_h$. Then, we extract a minimal subset $\mathcal{A}_h \subset \mathcal{T}_h$ such that

$$\sum_{K \in \mathcal{A}_h} \eta_K^2 \geq \theta_\mathrm{r} \sum_{K \in \mathcal{T}_h} \eta_K^2. \tag{4.4.1}$$

To obtain the refined mesh, each element in the subset $\mathcal{A}_h$ is bisected twice in 2D and three times in 3D. Similarly, to determine a set of cells to be coarsened, we extract a maximal subset $\mathcal{B}_h \subset \mathcal{T}_h$ such that

$$\sum_{K \in \mathcal{B}_h} \eta_K^2 \leq \theta_\mathrm{c} \sum_{K \in \mathcal{T}_h} \eta_K^2. \tag{4.4.2}$$

In general, we construct a fine mesh through successive bisections of a coarse mesh. An element in $\mathcal{B}_h$ is coarsened only when all of its neighboring fine elements, more precisely, the children of an element one refinement level above, are marked for coarsening. In general, we utilize for the triangulation the 'maximum smoothing' flag of `deal.II` [18]. For further details on local refinement and coarsening techniques for triangulations, see, for instance, [27].

4.5 Adaptive time-stepping and order control

In this section, we describe the procedure for adapting the order of the VSVO algorithm. We denote the VSVO algorithm with variable order ranging from k_min to k_max with $1 \leq k_\mathrm{min} \leq k_\mathrm{max} \leq 4$ by BDF(k_min - k_max). Let $\mathrm{TOL}_\mathrm{t} > 0$ be a given time tolerance and

$$s \coloneqq [1.44, 1.8, 1.6, 1.0] \tag{4.5.1}$$

be given scaling constants. We start by describing the order adaption based on the approximation of the local truncation error via extrapolation as described in Section 4.1.2. Let $\boldsymbol{m}_h^n$ denote the solution computed at t_n with step size τ_n. For each $k_\mathrm{min} \leq k \leq k_\mathrm{max}$, we compute

the predicted solution

$$\widehat{\boldsymbol{m}}_h^{n,(k)} \coloneqq \boldsymbol{P}^k(t_n),$$

where $\boldsymbol{P}^k$ denotes the polynomial predictor of order k, cf. (4.2.1). Then, the step-size scaling factor for order k can be computed by

$$\sigma^k = \left(\frac{s_k \mathrm{TOL}_\mathrm{t}}{\|\boldsymbol{m}_h^n - \widehat{\boldsymbol{m}}_h^{n,(k)}\|} \right)^{\frac{1}{k+1}} \tag{4.5.2}$$

with the scaling constant s_k provided in (4.5.1). Next, we compute the restricted time-step sizes

$$\tau_{n+1}^{(k)} = \max\left\{\tau_{\min},\, \underline{\kappa}_k \tau_n,\, \min\{\sigma^k \tau_n,\, \overline{\kappa}_k \tau_n,\, \tau_{\max}\}\right\}. \tag{4.5.3}$$

with the step-ratio restrictions $\underline{\kappa}_k$ and $\overline{\kappa}_k$ provided in (4.1.19). Alternatively, we may compute $\tau_{n+1}^{(k)}$ using the finite-difference approximations as described in (4.1.9). In total, this yields a set of potential step sizes

$$\tau_{n+1}^{(k_{\min})}, \ldots, \tau_{n+1}^{(k_{\max})}. \tag{4.5.4}$$

The order $k^\star$ for the next step, along with its associated step-size $\tau_{n+1}^{(k^\star)}$, is chosen as the one that maximizes the step-size by

$$k^\star = \arg \max_{k_{\min} \le k \le k_{\max}} \tau_{n+1}^{(k)}.$$

In general, we restrict the order change to ± 1 at each step. Once the order is changed, the new order $k^\star$ is maintained for $k^\star$ steps before considering another order change [43, 73].

STABILIZED ORDER CONTROL

In this section, we propose a stabilized order control designed to prevent frequent or redundant order changes. We follow several ideas from [82], which were applied to fluid-structure interaction problems in [83]. While the previous order control algorithm selects the order providing the largest time-step size according to (4.5.4), the stabilized version incorporates an additional stability test. When a method becomes unstable, the solution is often corrupted by high-frequency error components. These error modes are easier to detect by the higher-order derivatives of the numerical solution [82]. We consider the BDF(k) method for $k > 1$ as stable only if the predicted step-sizes for orders 1 to $k+1$ are strictly increasing, such that

$$\tau_{n+1}^{(1)} < \cdots < \tau_{n+1}^{(k)} < \tau_{n+1}^{(k+1)}. \tag{4.5.5}$$

Importantly, order k is only considered stable if the proposed step-size for order $k+1$, which is based on a higher-order derivative, exceeds that of order k. An increase in order from

k to $k+1$ is permitted only if order $k+1$ satisfies the stability condition (4.5.5) for $k+1$ consecutive steps. Following [82, Sct. 3.2.2], we maintain a constant order for 20 steps after any order change to allow the solution to settle. Furthermore, if an order is rejected and subsequently decreased, that specific order enters a *quarantine* period as described in [82, Sct. 3.2.3]. During this time, the order is unavailable for N steps, defined by

$$N = \max\Big\{50,\ 0.02\,\frac{T}{\tau}\max\{1,(k-2)^2\}\Big\},$$

where $T > 0$ denotes the final simulation time, k the prohibited order and τ the most recent time-step size. Consequently, the quarantine lasts for at least 50 steps and roughly for 2% of the total simulation time. Thus, if a method of order k is identified as unstable, all higher orders are also deemed unstable. In contrast to (4.1.19), we choose for the stabilized algorithm the step-size ratio restriction of $\underline{\kappa}_k = 0.9$ for $k = 1, 2, 3, 4$ and

$$\overline{\kappa}_1 = 4, \qquad \overline{\kappa}_2 = 1.6, \qquad \overline{\kappa}_3 = 1.3, \qquad \overline{\kappa}_4 = 1.1\,.$$

4.6 Adaptive algorithm

In this section, we describe a space and time adaptive algorithm we employ in this work for the LLG equation based on the tangent plane scheme (1.6.8). The algorithm combines the h-refinement described in Section 4.4 with the VSVO time-stepping described in Section 4.5. In general, we try to avoid a recomputation step, making use of the fact that the optimal time-step size usually does not change to much [56]. However, a time-step size rejection and recomputation may be easily incorporated. For a further discussion on the design of adaptive algorithms for parabolic equations we refer to [88].

In the following, we estimate the temporal error with the local truncation error of the BDF(k) method described in Section 4.1 and employ the step-size and order control following Section 4.5. For the spatial error estimation, we use either residual-based error estimators or gradient recovery techniques. More precisely, we define the local face residual error indicator by

$$\eta_K^{\Sigma}(\boldsymbol{m}_h^n) = \left(\sum_{E\subset\partial K} h_E \|R_E(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(E)}\right)^{1/2} \tag{4.6.1}$$

and the local cell residual error indicator by

$$\eta_K^{\mathcal{T}}(\boldsymbol{m}_h^n) = h_K \|R_K(\boldsymbol{m}_h^n)\|_{\boldsymbol{\mathcal{L}}^2(K)}. \tag{4.6.2}$$

Both indicators are motivated by the residual-based spatial error estimators introduced in

Definition 18. These local quantities provide the global error indicators

$$\eta^*(\boldsymbol{m}_h^n) = \left(\sum_{K \in \mathcal{T}_h^n} \left(\eta_K^\star(\boldsymbol{m}_h^n) \right)^2 \right)^{1/2} \tag{4.6.3}$$

for $\star \in \{\Sigma, \mathcal{T}, \boldsymbol{G}\}$, where $\eta_K^{\boldsymbol{G}}$ denotes the local gradient recovery error indicator defined in (4.3.1). With prescribed order bounds $1 \leq k_{\min} \leq k_{\max} \leq 4$, we describe now the space and time adaptive algorithm. Numerical results for the full adaptive algorithm are presented in Chapter 5.

Algorithm 59 (Space and time adaptivity for BDF($k_{\min}$ - $k_{\max}$))**.**

Input

- *A coarse initial mesh $\mathcal{T}_h^{ini}$, an initial magnetization $\boldsymbol{m}_h^0$, a final time $T > 0$.*
- *The tolerances* $\mathrm{TOL_s}$ *for space and* $\mathrm{TOL_t}$ *for time.*
- *The refinement and coarsening control parameters $\theta_{\mathrm{r}} > 0$ and $\theta_{\mathrm{c}} > 0$.*
- *A choice for the spatial error estimator $\star \in \{\Sigma, \mathcal{T}, \boldsymbol{G}\}$ in* (4.6.3)*.*

Precomputation

- *We start with the initial mesh $\mathcal{T}_h^{ini}$ and iteratively refine it to obtain $\mathcal{T}_h^0$. We use the spatial estimator $\eta^\star(\boldsymbol{m}_h^0)$ given in* (4.6.3) *to refine the elements in $\mathcal{A}_h$ as described in Section 4.4 using the parameters $\theta_{\mathrm{r}} = 0.3$ and $\theta_{\mathrm{c}} = 0.0$. We continue until the spatial estimator is below the prescribed tolerance* $\mathrm{TOL_s}$*.*
- *We compute $\boldsymbol{v}_h^0$ by* (2.3.6) *and estimate the preliminary initial time-step size $\widetilde{\tau}_1$ by*

$$\widetilde{\tau}_1 = \frac{\mathrm{TOL_t}}{\|\boldsymbol{v}_h^0\|}.$$

Then, using $\widetilde{\tau}_1$ we solve once (1.6.8) *with* BDF(1) *and the predictor*

$$\widehat{\boldsymbol{m}}_h^1 = \boldsymbol{m}_h^0 + \widetilde{\tau}_1 \boldsymbol{v}_h^0.$$

This provides us with the temporary values $\widetilde{\boldsymbol{m}}_h^1$ and $\widetilde{\boldsymbol{v}}_h^1$. We use them to compute the actual first time-step size τ_1 by

$$\tau_1 = c_s \left(\frac{2\,\mathrm{TOL}}{|\bar{\partial}_\tau^1 \boldsymbol{v}(t_1)|} \right)^{1/2}$$

according to (4.1.9)*, where we chose the safety constant $c_s = 1/2$.*

- *With the first time-step size τ_1, we recompute the first time-step using the* BDF(1) *method with the predictor $\widehat{\boldsymbol{m}}_h^1 = \widehat{\boldsymbol{m}}_h^0 + \tau_1 \boldsymbol{v}_h^0$. Thus, we end up with $\boldsymbol{m}_h^j$ and $\boldsymbol{v}_h^j$ for $j = 0, 1$. If we estimate the error using the predictor-based approach* (4.1.14)*, we perform two such steps since we require at least $\boldsymbol{m}_h^0$ and $\boldsymbol{m}_h^1$ to estimate the local truncation error for the* BDF(1) *method.*

- *If $k_{\min} > 1$, we gradually increase the BDF order each step by one, until we are at $k_{\min}$.*

Time-stepping for $n > k_{\min}$

- *We compute the predictor $\widehat{\boldsymbol{m}}_h^n$ by extrapolation from previous steps as described in Section 4.2.*
- *Optionally, we compute the stray field explicitly using $\boldsymbol{m}_h^{n-1}$ as described later in Section 5.8.*
- *Next, we compute the right-hand side $\boldsymbol{f}_h^n$.*
- *We then compute $\boldsymbol{v}_h^n$ by the tangent plane scheme* (1.6.8) *and update the magnetization $\boldsymbol{m}_h^n$ by* (1.6.5).
- *We continue by normalizing the magnetization $\widetilde{\boldsymbol{m}}_h^n \coloneqq \boldsymbol{m}_h^n/|\boldsymbol{m}_h^n|$ prior to evaluating the spatial error estimator. This improves the effectiveness of coarsening, cf. Remark 61. We emphasize that the normalized magnetization is used only for the error computation and is not propagated to the next step. Next, we adapt the spatial mesh $\mathcal{T}_h^n$ using the estimator $\eta^\star(\widetilde{\boldsymbol{m}}_h^n)$ given in* (4.6.3) *with the marking strategy described in Section 4.4 to obtain the spatial mesh $\mathcal{T}_h^{n+1}$. Refinement is only performed if the global estimated error is larger than the prescribed space tolerance* $\mathrm{TOL_s}$. *Unless stated otherwise, we employ the marking parameters $\theta_{\mathrm{c}} = 0.1$ and $\theta_{\mathrm{r}} = 0.15$. Finally, we interpolate all quantities required for subsequent computations onto the updated grid $\mathcal{T}_h^{n+1}$.*
- *We selected the next order $k^\star$ and time-step size $\tau_{n+1}^{(k^\star)}$ by equilibrating the local truncation error as described in Section 4.5. If $t_n + \tau_{n+1}^{(k^\star)} > T$, we adjust $\tau_{n+1}^{(k^\star)}$ such that $t_n + \tau_{n+1}^{(k^\star)} = T$.*

The following remark points out that the start-up in Algorithm 59 may not be directly applicable to uniform BDF schemes of higher order.

Remark 60 (Algorithm start-up)**.** *The initialization of Algorithm 59 is not necessarily appropriate for higher order uniform BDF schemes. For instance, we observed for the sequence* BDF(1) $\to$ BDF(2) $\to$ BDF(3) *a reduction of the convergence rate (expected $\mathcal{O}(\tau^3)$ and obtained only $\mathcal{O}(\tau^2)$). To prevent such order reduction, one may initialize the multistep method using higher-order schemes, for example, an L-stable singly diagonal implicit Runge–Kutta (SDIRK) method. Nevertheless, for Algorithm 59, we did not encounter any issues starting the adaptive algorithm with the* BDF(1) *method using a sufficiently small time-step size τ_1 since the aim is to achieve a prescribed local error tolerance.*

We conclude this section with a remark regarding the use of normalization as a post-processing step before performing spatial error estimation.

Remark 61 (Normalization as post-processing)**.** *For the purpose of spatial error estimation, we normalize the magnetization at each degree of freedom to enhance the effectiveness of the coarsening procedure, as demonstrated later in Section 5.5.2. This normalization is employed solely for error estimation and is not incorporated into the time-stepping algorithm itself. When applied only as a post-processing step, the normalized magnetization satisfies the same approximation properties. This was established in [30] for the harmonic map heat flow with the* BDF(1) *method.*

CHAPTER 5

Numerical results

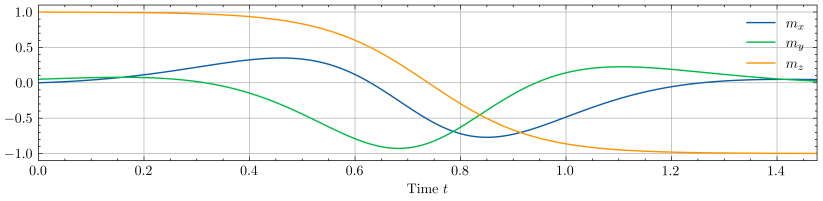
In this section, we apply the space and time adaptive algorithm described in Section 4.6 to several numerical examples. We begin with an ordinary differential equation example for the LLG equation to evaluate the performance of different time adaptive strategies. In particular, we take small values for the damping coefficient α into account. Subsequently, we demonstrate that all a posteriori error estimators introduced in Definition 18 converge with the optimal order with respect to either the time-step size or the spatial mesh size. We then apply the fully adaptive algorithm to a variety of examples, including a moving bump, a domain wall and the physically relevant μMAG #4 example. To simplify the notation, we abbreviate the following norms by

$$\|\cdot\|_{\mathcal{L}^\infty(\boldsymbol{\mathcal{L}}^2)} = \|\cdot\|_{\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{L}}^2(\Omega))},$$
$$\|\cdot\|_{\mathcal{L}^\infty(\mathcal{L}^\infty)} = \|\cdot\|_{\mathcal{L}^\infty(0,T;\mathcal{L}^\infty(\Omega))}.$$

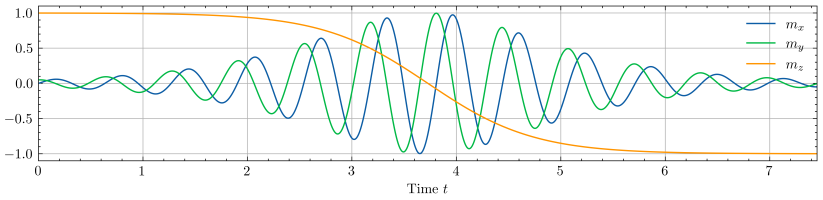
The final simulation time T is explicitly given in the respective examples.

5.1 Implementation details

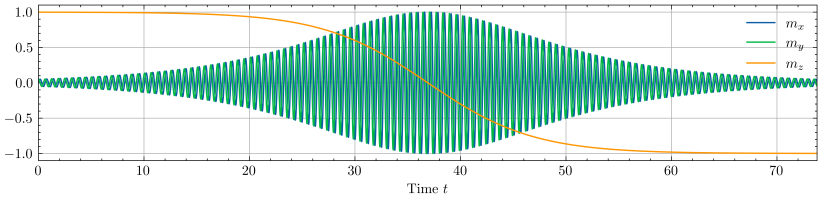
In general, we compute the numerical results using the open-source finite element library `deal.II` version `9.5.1` implemented in `C++` [18]. The `deal.II` framework offers essential features for performing space adaptive simulations, enabling mesh refinement and coarsening according to a fixed fraction of the estimated error as defined in (4.4.1) and (4.4.2). The `deal.II` framework is based on quadrilaterals in 2D and hexahedrons in 3D. We solve the resulting linear systems using the UMFPACK (Unsymmetric MultiFrontal PACKage) package [50] for the LU-decomposition or the GMRES (generalized minimal residual) iterative solver with an incomplete LU-decomposition preconditioner [108, 109].



(a) $\alpha = 1.0$



(b) $\alpha = 0.1$



(c) $\alpha = 0.01$

Figure 5.2.1: Evolution of the x-, y- and z-component for the solution of (5.2.1) with $\alpha = 1, 0.1, 0.01$. The numerical solution was computed using the uniform time-step size BDF(2) method with $\tau = 0.001$.

5.2 Adaptive time-stepping for a LLG ODE

In this section, we study the example from [114, Sct. 7.4] and consider the LLG equation (1.6.1) in a spatially uniform form with the effective field $\boldsymbol{h}_{\text{eff}} = \boldsymbol{f}$, which reduces to the ordinary differential equation (ODE)

$$\begin{aligned} \alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} &= \boldsymbol{P}(\boldsymbol{m}) \boldsymbol{f}, \\ \partial_t \boldsymbol{m} \cdot \boldsymbol{m} &= 0 \end{aligned}$$

for $t \in (0, T)$, where $\boldsymbol{P}(\boldsymbol{m}) = \boldsymbol{I} - \boldsymbol{m}\boldsymbol{m}^\top$ denotes the orthogonal projection onto the tangent plane. In the following, we compare adaptive time-stepping strategies based on the local error estimators introduced in Section 4.1. We consider an initial magnetization that is close to $[0, 0, 1]^\top$ and apply a constant external field of $\boldsymbol{f} = [0, 0, H]^\top$ with $H = -10$ to reverse the orientation of the magnetization in the direction of the external field. To avoid starting from an equilibrium state, we slightly perturb the initial magnetization and employ the initial data

$$\boldsymbol{m}_0 = \begin{bmatrix} m_x^0 \\ m_y^0 \\ m_z^0 \end{bmatrix} = \frac{1}{\sqrt{\delta^2 + 1}} \begin{bmatrix} 0 \\ \delta \\ 1 \end{bmatrix}$$

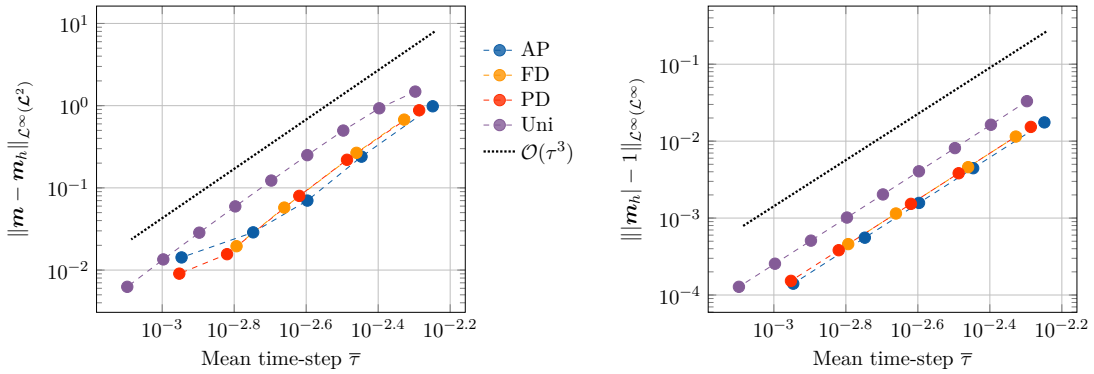


Figure 5.2.2: The maximal $\mathcal{L}^2$-error (left) and the maximal normalization error (right) for the solution of (5.2.1) with $\alpha = 0.01$ and $T = 80$. The results are computed with the uniform and variable step-size BDF(2) method. For the adaptive case, the time-steps are selected by the finite-difference (FD) approximation (4.1.9), the predictor-based strategy (PD) (4.1.14) and the a posteriori estimator (AP) (4.1.17) with $i = 2$. Purple lines represent the uniform step-size BDF(2) method (Uni). The horizontal axis denotes the average time-step size $\overline{\tau} = T/N_\tau$, where N_τ is the total number of time-steps, while the initial time-step size for the adaptive methods is fixed at $\tau_0 = 10^{-3}$.

with the perturbation $\delta = 0.05$. For a physical motivation, we refer to [3, 114]. Defining $C = \operatorname{arctanh}(m_z^0)$ and $\omega = H/(1+\alpha^2)$, we can express the exact solution by

$$\boldsymbol{m}(t) = \boldsymbol{m}(t, \boldsymbol{x}) = \begin{bmatrix} A(t)\big(m_x^0 \cos(\omega t) - m_y^0 \sin(\omega t)\big) \\ A(t)\big(m_x^0 \sin(\omega t) + m_y^0 \cos(\omega t)\big) \\ \tanh(\alpha\omega t + C) \end{bmatrix}, \tag{5.2.1}$$

where A is defined by

$$A(t) \coloneqq \frac{\sqrt{1 - m_z(t)^2}}{\sqrt{1 - (m_z^0)^2}}.$$

Exemplary results for the Gilbert damping parameter $\alpha = 1, 0.1, 0.01$ are illustrated in Figure 5.2.1. Smaller values of α lead to increasingly frequent oscillations in the magnetization dynamics. We measure the error in the $\mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{L}}^2(\Omega))$-norm by

$$\|\boldsymbol{m} - \boldsymbol{m}_h\|_{\mathcal{L}^\infty(\boldsymbol{\mathcal{L}}^2)}, \tag{5.2.2}$$

where $\boldsymbol{m}_h$ is the linear reconstruction defined in (2.1.4). Additionally, we consider the maximal unit-length constraint error

$$\||\boldsymbol{m}_h| - 1\|_{\mathcal{L}^\infty(\mathcal{L}^\infty)} \tag{5.2.3}$$

since we do not employ an explicit normalization step.

First, we begin by comparing the various time-step size selection strategies introduced

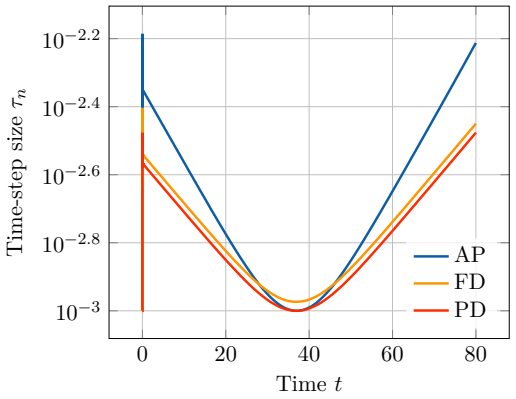


Figure 5.2.3: Time-step size selection strategies for the variable step-size BDF(2) method, including the finite-difference (FD) approximation (4.1.9), the predictor-based strategy (PD) (4.1.14) and the a posteriori estimator (AP) (4.1.17) for $i = 2$. The initial time-step size is fixed to $\tau_0 = 10^{-3}$.

earlier. We set the final simulation time $T = 80$ and choose $\alpha = 0.01$. Since we do not specify an initial time-step selection for the a posteriori estimator (AP) (4.1.17), we set $\tau_0 = 10^{-3}$ to improve comparability between the different strategies. In Figure 5.2.2, we compute the error between the exact solution $\boldsymbol{m}$ given in (5.2.1) and the linear reconstruction $\boldsymbol{m}_h$ defined in (2.1.4). Although we employ the second-order BDF(2) method, we observe third-order convergence for the maximal $\boldsymbol{\mathcal{L}}^2$-error (5.2.2) and the maximal normalization error (5.2.3) with respect to the mean time-step size $\overline{\tau} = T/N_\tau$, where N_τ denotes the total number of time-steps. Furthermore, the adaptive simulations with the finite-difference (FD) approximation (4.1.9), the predictor-based strategy (PD) (4.1.14) and the AP estimator (4.1.17) consistently yield lower error values for the maximal $\boldsymbol{\mathcal{L}}^2$-error (5.2.2) and the maximal normalization error (5.2.3) in comparison to the uniform step-size simulation. As illustrated in Figure 5.2.3, each of the time-step size strategies considered leads to a smoothly varying sequence of time-step sizes. Moreover, the time-step size is reduced while the magnetization's z-component changes sign.

Second, we consider the VSVO BDF(1 - 4) algorithm as described in Section 4.5. We exclude the AP estimator in this context since it is specifically designed for the fixed-order BDF(2) method. The local truncation error is estimated using either the FD or PD approximation, as described in Section 4.5. In Figure 5.2.4, we observe convergence of order 5 for the VSVO BDF(1 - 4) algorithm for the maximal $\mathcal{L}^2$-error (5.2.2) and the maximal $\mathcal{L}^\infty$-normalization error (5.2.3) with respect to the mean time-step size $\overline{\tau}$.

In Figure 5.2.5, we compare the evolution of the time-step size and the BDF orders selected by the FD and PD approaches. We observe that the changes in the time-step sizes and the orders follow a very similar pattern for both methods. Overall, the choice of the error estimation method has a negligible impact on the general performance of the VSVO BDF(1 - 4) algorithm.

Next, we briefly discuss the implications of a small damping parameter α as employed in this section. The eigenvalues of the linearized LLG operator are determined by the precessional

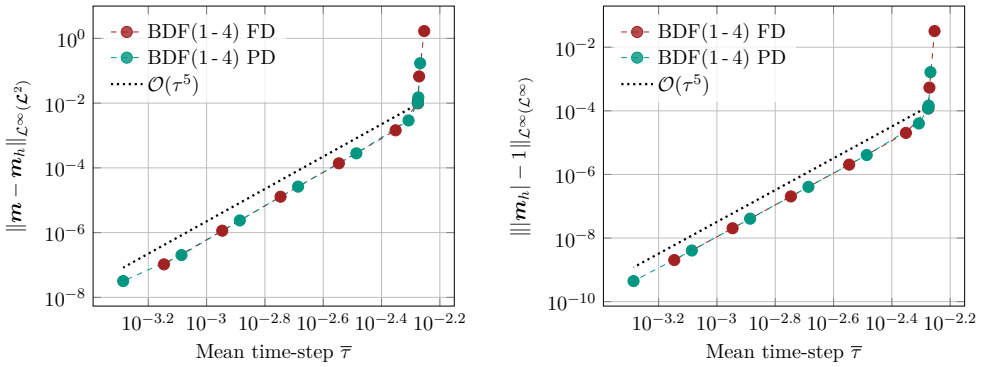


Figure 5.2.4: The maximal $\mathcal{L}^2$-error (left) and the maximal normalization error (right) for the solution of (5.2.1) with $\alpha = 0.01$ and $T = 80$. The results are computed with the VSVO BDF(1-4) method as described in Section 4.5. The horizontal axis denotes the average time-step size $\overline{\tau} = T/N_\tau$, where N_τ is the total number of time-steps. The corresponding time tolerance ranges from 10^{-4} to 10^{-12}.

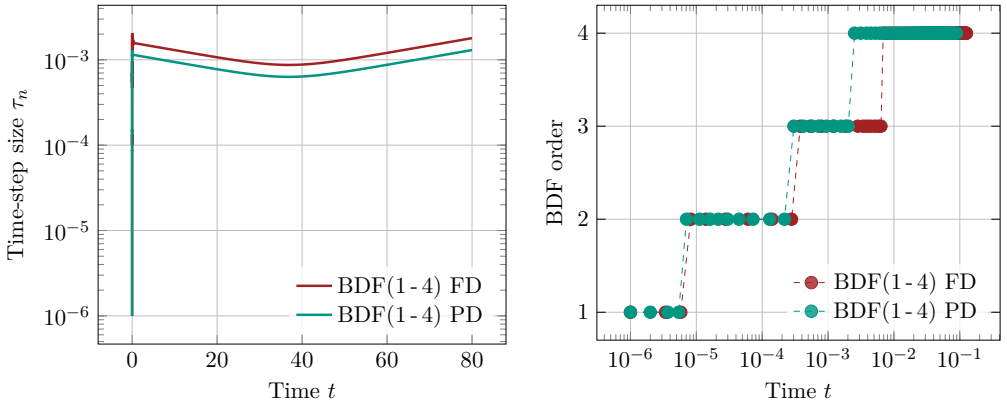


Figure 5.2.5: Evolution of the time-step size (left) and the selected BDF order (right) for the solution of (5.2.1) with $\alpha = 0.01$ and $T = 80$ for the VSVO BDF(1-4) method as described in Section 4.5 with $\mathrm{TOL_t} = 10^{-11}$. On the right only the first 100 time-steps are shown. Afterwards the BDF order remains at 4.

and dissipative dynamics. As highlighted in Example 39, as the damping coefficient α tends to 0, the eigenvalues approach the imaginary axis. Although higher-order BDF methods generally lose stability when the spectrum lies near the imaginary axis, cf. Figure 3.2.2, we observe no stability issues for this example, although we have $\alpha = 0.01$. A further discussion on the performance of higher-order BDF schemes in the presence of nearly imaginary eigenvalues can be found in [83, Sct. 2.5] for fluid-structure interaction problems. We emphasize that the VSVO algorithm is designed to mitigate potential stability issues by reverting to A-stable methods, such as the BDF(1) or the uniform step-size BDF(2) scheme.

For finely discretized spatial domains, the computational cost of time error estimation is dominated by the evaluation of the $\mathcal{L}^2$-norm, either for the difference between the predictor and the solution or for the higher-order time derivative. As a result, the FD and PD approaches require essentially the same total computation time since each estimate requires a single $\mathcal{L}^2$-norm evaluation. Overall, the cost of the temporal estimator within the full adaptive

algorithm remains negligible, as discussed later in Section 5.5. In the following examples, we stick to the PD local truncation error estimation for the VSVO BDF(1-4) algorithm.

5.3 Residual-based a posteriori error estimators

In this section, we present numerical experiments for the a posteriori error estimators introduced in Definition 18, which lead to an a posteriori error bound as stated in Theorem 19. Furthermore, we demonstrate that all a posteriori error estimators converge at the optimal rate with respect to either the uniform time-step size or the uniform mesh size. We abbreviate the $\boldsymbol{\mathcal{H}}^1$-error, estimated in Theorem 19, by

$$\operatorname{err}_{\boldsymbol{\mathcal{H}}^1}^n \coloneqq \|\boldsymbol{\nabla}(\boldsymbol{m}(t_n) - \boldsymbol{m}_h^n)\|, \qquad \operatorname{err}_{\boldsymbol{\mathcal{H}}^1} \coloneqq \|\boldsymbol{\nabla}(\boldsymbol{m} - \boldsymbol{m}_h)\|_{\mathcal{L}^\infty(0,t_N;\boldsymbol{\mathcal{L}}^2(\Omega))}. \tag{5.3.1}$$

For the implementation, we evaluate the projection error estimator $\mathcal{P}$ using Simpsons rule on each interval, which yields

$$\mathcal{P} = \int_0^{t_N} \|\boldsymbol{q}_h(t)\|_{\boldsymbol{\mathcal{H}}^1}^2 \,\mathrm{d}t \approx \sum_{n=2}^{N} \frac{\tau}{6}\Big(\|\boldsymbol{q}_h(t_{n-1})\|_{\boldsymbol{\mathcal{H}}^1}^2 + 4\|\boldsymbol{q}_h(t_{n-1/2})\|_{\boldsymbol{\mathcal{H}}^1}^2 + \|\boldsymbol{q}_h(t_n)\|_{\boldsymbol{\mathcal{H}}^1}^2\Big),$$

where $t_{n-\frac{1}{2}} = t_{n-1} + \tau/2 = t_n - \tau/2$ and

$$\boldsymbol{q}_h(t) \coloneqq \partial_t \boldsymbol{M}_h(t) - \boldsymbol{P}(\boldsymbol{M}_h(t))\partial_t \boldsymbol{M}_h(t).$$

Recall that $\boldsymbol{P}(\boldsymbol{m}) = \boldsymbol{I} - \boldsymbol{m}\boldsymbol{m}^\top$ denotes the orthogonal projection onto the tangent plane. We abbreviate the sums of the combined a posteriori error estimator components by

$$\begin{aligned} \mathcal{E} &\coloneqq \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4 + \mathcal{E}_5 + \mathcal{E}_6, & \mathcal{C} &\coloneqq \mathcal{C}_1 + \mathcal{C}_2, \\ \Lambda &\coloneqq \Lambda_1 + \Lambda_2 + \Lambda_3, & \mathcal{Q} &\coloneqq \mathcal{Q}_1 + \mathcal{Q}_2 \end{aligned} \tag{5.3.2}$$

where the estimators are introduced in Definition 18.

We first demonstrate that the temporally dominant error estimators converge with the expected experimental order of convergence. Similarly, the spatially dominant estimators are shown to converge at the expected rates. We effectively obtain a split of the a posteriori error estimators into separate spatial and temporal contributions. We further investigate the spatial error estimator Λ by highlighting the distinctions between the element and face residual components.

5.3.1 Convergence for uniform time-steps

In this section, we study the example from [7, Eq. (9.2)], for which the temporal discretization error dominates the spatial discretization error. We consider the domain $\Omega = (0,1) \times (0,1)$ and choose the damping parameter $\alpha = 0.2$ and the exchange length $\ell_{\mathrm{ex}}^2 = 1$. We fix the final simulation time $T = 0.1$. For the spatial discretization, we use a uniform mesh

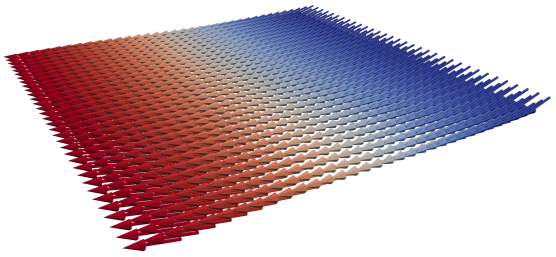


Figure 5.3.1: Initial data of the example (5.3.3). The color represents the z-component of the magnetization, ranging from $-1/4$ to $1/4$ (blue to red).

consisting of 128×128 quadrilaterals together with quadratic finite elements. For the time discretization, we employ the uniform BDF(2) method with varying uniform time-step size τ. We consider the tangent plane scheme as described in Section 2.3. We prescribe the exact solution $\boldsymbol{m} : [0,T] \times \Omega \to \mathbb{R}^3$ by

$$\boldsymbol{m}(t,\boldsymbol{x}) \coloneqq \begin{bmatrix} -h(x_1)\sin(3\pi t/T) \\ \sqrt{1-h(x_1)^2} \\ -h(x_1)\cos(3\pi t/T) \end{bmatrix} \tag{5.3.3}$$

for $(t,\boldsymbol{x}) \in [0,T] \times \Omega$. Here, the function h is defined by

$$h(x) \coloneqq x^3 - \frac{3x^2}{2} + \frac{1}{4}$$

for $x \in \mathbb{R}$. The initial data is depicted in Figure 5.3.1. By construction, we have $|\boldsymbol{m}(t,x)| = 1$ for $(t,\boldsymbol{x}) \in [0,T] \times \Omega$. Finally, we compute the right-hand side $\boldsymbol{f}$ such that $\boldsymbol{m}$ defined in (5.3.3) is the exact solution of the alternative form (1.5.1) of the LLG equation, that is,

$$\boldsymbol{f} = \alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} - \ell_{\mathrm{ex}}^2 \Delta \boldsymbol{m},$$

where we can neglect the orthogonal projection onto the tangent plane since we are solving in the tangent space anyway.

As illustrated in Figure 5.3.2, the combined temporal error estimator $\mathcal{E}$ defined in (5.3.2) as well as the maximum gradient error (5.3.1) converges with second order. Moreover, the projection error estimator $\mathcal{P}$ achieves a convergence of order three, while the combined extrapolation error estimator $\mathcal{Q}$ converges with fourth-order accuracy. In contrast, the combined spatial error estimator Λ is, as expected, independent of the temporal discretization. As demonstrated later in Figure 5.3.5, the spatial error estimator Λ converges with respect to the uniform spatial mesh size. Similarly, the finite element space conforming estimator $\mathcal{C}$ remains (almost) constant but converges with respect to the spatial mesh size as demonstrated later in Table 5.3.1. In total, the sum of the a posteriori error estimators (which corresponds effectively to the temporal error estimator $\mathcal{E}$) yields a reliable upper bound for the maximum gradient error, as stated in Theorem 19.

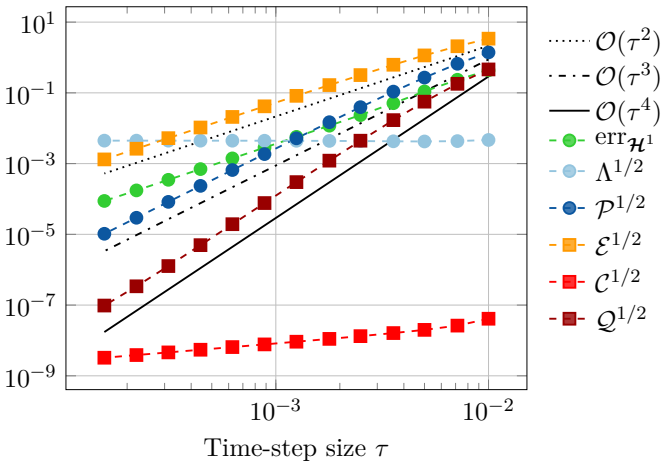


Figure 5.3.2: The space error estimator Λ, the time error estimator $\mathcal{E}$, the projection error estimator $\mathcal{P}$, the finite element space conforming estimator $\mathcal{C}$ and the extrapolation error estimator $\mathcal{Q}$ defined in Definition 18 and the maximum gradient error $\mathrm{err}_{\mathcal{H}^1}$ (5.3.1) with respect to the uniform time-step size τ for example (5.3.3) using the BDF(2) method. The sum of all a posteriori error estimators corresponds effectively to the time error estimator $\mathcal{E}$ (yellow).

The individual components of the combined time error estimator with respect to the time-step size τ are shown in Figure 5.3.3. While the time error estimator $\mathcal{E}_3$ exhibits convergence of order 4, all remaining time error estimators converge with second order accuracy.

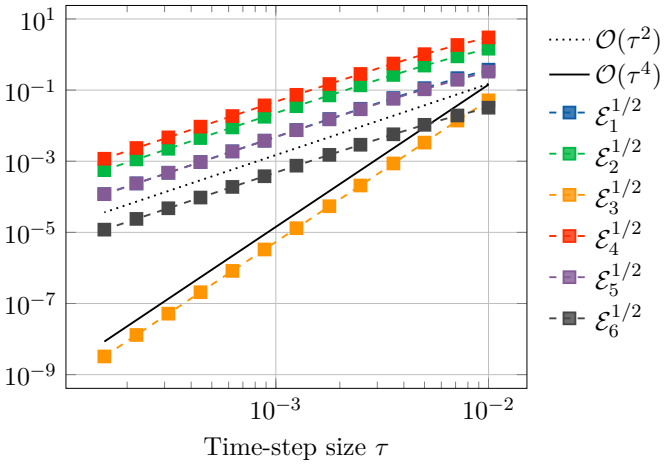


Figure 5.3.3: The time and reconstruction a posteriori error estimators $\mathcal{E}_1, \ldots, \mathcal{E}_6$ defined in Definition 18 with respect to the uniform time-step size τ for example (5.3.3) using the BDF(2) method.

5.3.2 CONVERGENCE FOR UNIFORM MESH SIZES

In this section, we study the example from [7, Eq. (9.1)], which is constructed in such a way that the spatial error dominates. We consider the domain $\Omega = (0,1) \times (0,1)$ and choose the damping parameter $\alpha = 0.2$ and the exchange length $\ell_{\mathrm{ex}}^2 = 1$. Let $C = 400$, fix the

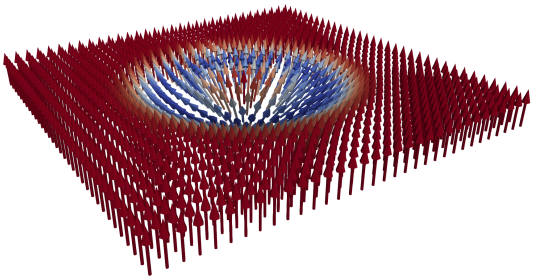


Figure 5.3.4: Initial data of the example (5.3.5). The color represents the z-component of the magnetization.

final time $T = 0.01$ and set $T_0 = 0.06$. For the time discretization, we employ the uniform step-size BDF(2) method with time-step size $\tau = T/1000$. We consider the discretization as in Section 2.3 for globally continuous piecewise polynomials of degree $p = 1, 2, 3, 4$. Additionally, we define the squared distance from the center of the domain Ω by

$$d(\boldsymbol{x}) \coloneqq (x_1 - 1/2)^2 + (x_2 - 1/2)^2 \tag{5.3.4}$$

for $\boldsymbol{x} = (x_1, x_2)^\top \in \mathbb{R}^2$ and introduce the time-dependent function

$$g(t) \coloneqq \frac{T_0 + 0.1}{T_0 + 0.1 - t}$$

for $t \in [0, T_0]$. We prescribe the exact solution $\boldsymbol{m} : [0, T] \times \Omega \to \mathbb{R}^3$ by

$$\boldsymbol{m}(t, \boldsymbol{x}) \coloneqq \begin{bmatrix} C \exp\left(\frac{-g(t)}{1/4 - d(\boldsymbol{x})}\right)(x_1 - 1/2) \\ C \exp\left(\frac{-g(t)}{1/4 - d(\boldsymbol{x})}\right)(x_2 - 1/2) \\ \sqrt{1 - C^2 \exp\left(\frac{-2g(t)}{1/4 - d(\boldsymbol{x})}\right) d(\boldsymbol{x})} \end{bmatrix} \tag{5.3.5}$$

if $d(\boldsymbol{x}) \leq 1/4$ and otherwise by $\boldsymbol{m}(\boldsymbol{x}, t) \coloneqq [0, 0, 1]^\top$ for $(t, \boldsymbol{x}) \in [0, T] \times \Omega$. The initial data is depicted in Figure 5.3.4. Obviously, we have $|\boldsymbol{m}(t, x)| = 1$ for $(t, \boldsymbol{x}) \in [0, T] \times \Omega$. Finally, we compute the right-hand side $\boldsymbol{f}$ such that $\boldsymbol{m}$ defined in (5.3.5) is the exact solution of the alternative form (1.5.1) of the LLG equation, that is,

$$\boldsymbol{f} = \alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} - \ell_{\mathrm{ex}}^2 \Delta \boldsymbol{m},$$

where we can neglect the orthogonal projection onto the tangent plane since we are solving in the tangent space anyway.

As illustrated in Figure 5.3.5, both the combined spatial error estimator Λ defined in (5.3.2) and the projection error estimator $\mathcal{P}$ given in Definition 18 converge with the same order as the maximum gradient error (5.3.1) for all polynomial degrees $p = 1, 2, 3, 4$. Moreover, Figure 5.3.5 indicates that the estimators yield an upper bound for the maximum gradient error, as stated in Theorem 19. Similarly, Table 5.3.6 shows that the finite element space conforming error

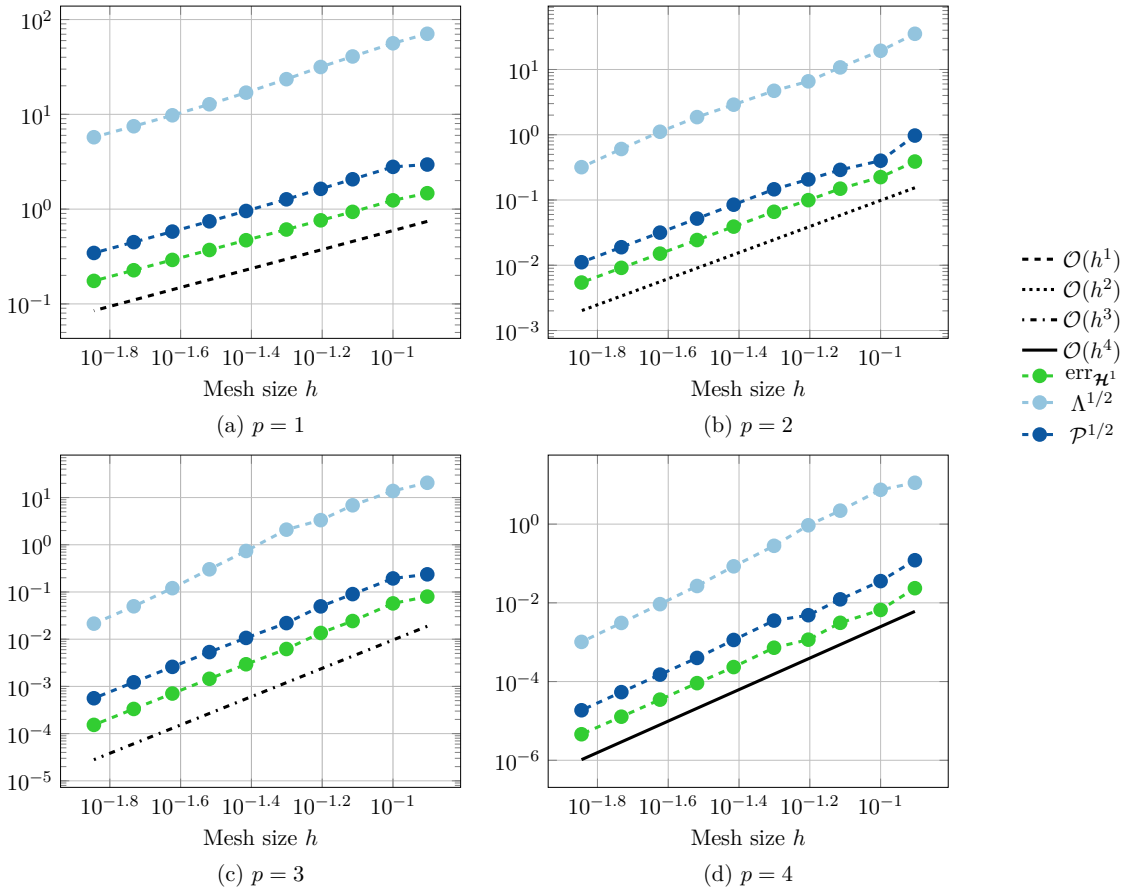
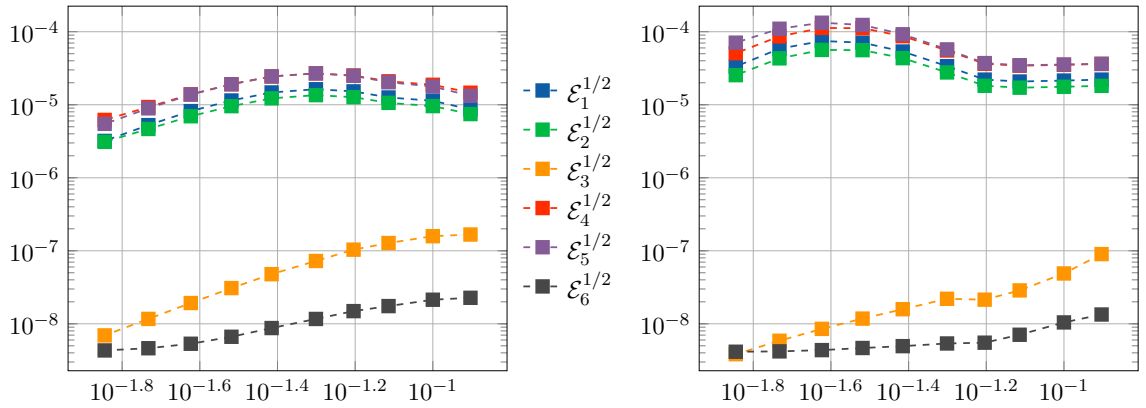


Figure 5.3.5: The space and projection a posteriori error estimator Λ and $\mathcal{P}$ defined in Definition 18 as well as the maximum gradient error $\mathrm{err}_{\mathcal{H}^1}$ (5.3.1) with respect to the uniform spatial mesh size h for example (5.3.5) and polynomial degrees $p = 1, 2, 3, 4$.

	$p = 1$		$p = 2$		$p = 3$		$p = 4$	
h	Value	Rate	Value	Rate	Value	Rate	Value	Rate
1/26	3.802e-06	—	5.027e-06	—	3.110e-07	—	6.616e-08	—
1/33	1.543e-06	3.791	3.047e-06	2.099	1.153e-07	4.168	2.705e-08	3.753
1/42	6.657e-07	3.719	1.747e-06	2.457	3.442e-08	5.341	1.050e-08	4.175
1/54	2.957e-07	3.226	9.337e-07	2.490	9.613e-09	5.071	3.704e-09	4.148
1/70	1.324e-07	3.090	4.693e-07	2.648	2.573e-09	5.093	1.194e-09	4.356

Table 5.3.1: The combined finite element space conforming error estimators $\mathcal{C}$ (5.3.2) introduced in Definition 18 with respect to the uniform spatial mesh size h for example (5.3.5) for the polynomial degrees $p = 1, 2, 3, 4$.

estimator $\mathcal{C}$ converges at least with order p for all finite element degrees $p = 1, 2, 3, 4$. The individual components of the combined time error estimator with respect to the spatial mesh size h are shown in Figure 5.3.6. As expected, the time error estimators remain (almost) constant and are thus mostly independent of the spatial discretization. As illustrated earlier in Figure 5.3.3, the time error estimator converge with respect to the time-step size. While the error estimators $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_4, \mathcal{E}_5$ exhibit similar behavior, the components $\mathcal{E}_3, \mathcal{E}_6$ are several orders of magnitude smaller.



(a) $p = 1$ (b) $p = 2$

Figure 5.3.6: The time and reconstruction a posteriori error estimators $\mathcal{E}_1, \dots, \mathcal{E}_6$ defined in Definition 18 with respect to the uniform spatial mesh size h for example (5.3.5) for the polynomial degrees $p = 1, 2$.

Contributions of the cell and face residuals to the spatial error estimators

Recall the space error estimator

$$\Lambda_1 = \max_{0\le n\le N} \eta(\boldsymbol{m}_h^n; \mathcal{H}^1) = \max_{0\le n\le N} \left(\sum_{K\in\mathcal{T}_h^n} h_K^2 \|R_K(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(K)} + \sum_{E\in\Sigma_n} h_E \|R_E(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(E)} \right)^{1/2} \tag{5.3.6}$$

introduced in Definition 18, where η is defined in (2.2.7). In order to investigate the respective contributions of the cell and face residuals to the spatial error estimator Λ_1, we consider both terms individually, namely, the face residual estimator

$$\eta^{\Sigma} = \max_{0\le n\le N} \left(\sum_{E\in\Sigma_n} h_E \|R_E(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(E)} \right)^{1/2} \tag{5.3.7}$$

and the cell residual estimator

$$\eta^{\mathcal{T}} = \max_{0\le n\le N} \left(\sum_{K\in\mathcal{T}_h^n} h_K^2 \|R_K(\boldsymbol{m}_h^n)\|^2_{\boldsymbol{\mathcal{L}}^2(K)} \right)^{1/2}, \tag{5.3.8}$$

cf. (4.6.3). Thus, by (5.3.6) we have that $\max_{0\le n\le N} \left(\eta(\boldsymbol{m}_h^n; \mathcal{H}^1)\right)^2 = \left(\eta^{\Sigma}\right)^2 + \left(\eta^{\mathcal{T}}\right)^2$.

The cell and face residual estimators in a posteriori error estimation are known to depend on the polynomial degree of the finite element space. According to [119, Remark 1.17] and [42, 90], Yu showed in [122, 123] that for finite element discretizations on a uniform square grid, the error is controlled either by the cell residuals or the face residuals, depending on whether the polynomial degree of the finite element functions is even or odd. For linear finite

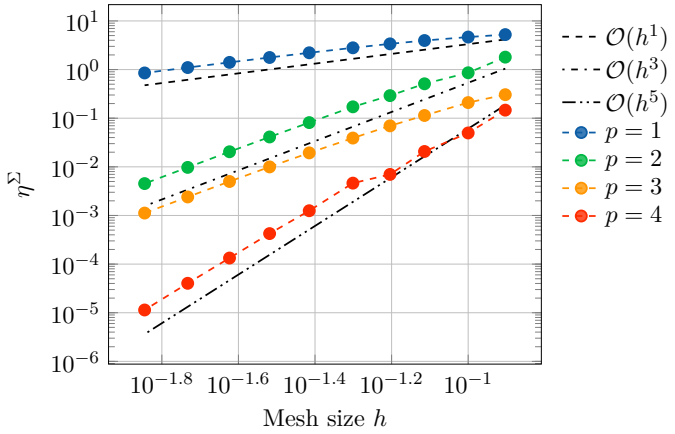


Figure 5.3.7: The face residual error estimator η^{Σ} of Λ_1 defined in (5.3.7) for example (5.3.5) and polynomial degrees $p = 1, 2, 3, 4$.

elements on a shape-regular triangular mesh, the face residuals dominate [42]. We observe this behavior in the numerical results depicted in Figure 5.3.7 and 5.3.8. As indicated in Figure 5.3.7, the face residual error estimator (5.3.7) exhibits a convergence order of p for odd polynomial degrees $p = 1, 3$, while for even degrees $p = 2, 4$, an increased convergence rate of order $p + 1$ is observed. In contrast, the cell residual error estimator (5.3.8) exhibits convergence of order p for all polynomial degrees $p = 1, 2, 3, 4$, see Figure 5.3.8.

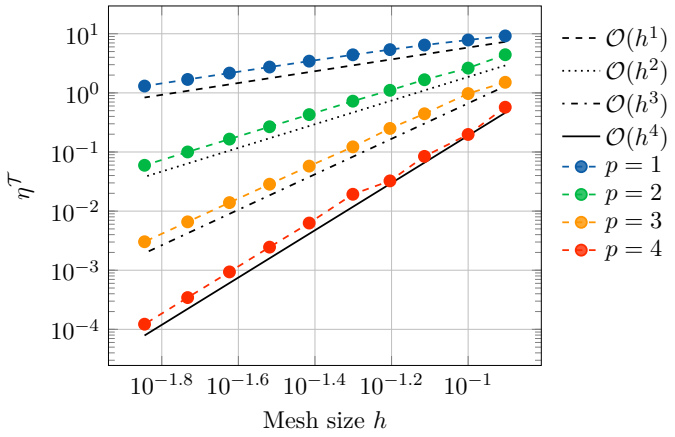


Figure 5.3.8: The cell residual error estimator $\eta^{\mathcal{T}}$ of Λ_1 defined in (5.3.8) for example (5.3.5) and polynomial degrees $p = 1, 2, 3, 4$.

5.4 COMPUTATIONAL EFFORT

In this section, we analyze the performance of the solver of the linear equation system based on the tangent plane scheme. To this end, we consider the space-dominant example described in Section 5.3.2 and compute 10 time-steps using a constant step size $\tau = 10^{-5}$. We begin by

recalling the fully discrete equation (1.6.8), which seeks $(\boldsymbol{v}_h^n, \lambda_h^n) \in \boldsymbol{V}_h^n \times V_h^n$ such that

$$\alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \ell_{\mathrm{ex}}^2\langle \boldsymbol{\nabla}\boldsymbol{m}_h^n, \boldsymbol{\nabla}\boldsymbol{\varphi}_h^n\rangle + \langle \widehat{\boldsymbol{m}}_h^n \cdot \boldsymbol{\varphi}_h^n, \lambda_h^n\rangle = \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n\rangle,$$
$$\langle \widehat{\boldsymbol{m}}_h^n \cdot \boldsymbol{v}_h^n, \psi_h^n\rangle = 0$$

for all $(\boldsymbol{\varphi}_h^n, \psi_h^n) \in \boldsymbol{V}_h^n \times V_h^n$. By substituting the relation $\tau_n \boldsymbol{v}_h^n = \sum_{j=0}^k \delta_j \boldsymbol{m}_h^{n-j}$ with $k \in \mathbb{N}$, cf. (1.6.5), we obtain the following linear system

$$\alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \ell_{\mathrm{ex}}^2 \frac{\tau_n}{\delta_0}\langle \boldsymbol{\nabla}\boldsymbol{v}_h^n, \boldsymbol{\nabla}\boldsymbol{\varphi}_h^n\rangle + \langle \widehat{\boldsymbol{m}}_h^n \cdot \boldsymbol{\varphi}_h^n, \lambda_h^n\rangle$$
$$= \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n\rangle + \ell_{\mathrm{ex}}^2 \sum_{j=1}^k \frac{\delta_j}{\delta_0}\langle \boldsymbol{\nabla}\boldsymbol{m}_h^{n-j}, \boldsymbol{\nabla}\boldsymbol{\varphi}_h^n\rangle,$$
$$\langle \widehat{\boldsymbol{m}}_h^n \cdot \boldsymbol{v}_h^n, \psi_h^n\rangle = 0$$

for all $(\boldsymbol{\varphi}_h^n, \psi_h^n) \in \boldsymbol{V}_h^n \times V_h^n$. Let $\{\phi_i\}_{i=1}^N$ denote the nodal basis of V_h^n with $N = \dim(V_h^n)$. To represent the system algebraically, we define the mass matrix $M \in \mathbb{R}^{N\times N}$ with entries $m_{ij} = \langle \phi_i, \phi_j\rangle$ and the stiffness matrix $S \in \mathbb{R}^{N\times N}$ with entries $s_{ij} = \langle \boldsymbol{\nabla}\phi_i, \boldsymbol{\nabla}\phi_j\rangle$. The terms involving the predictor $\widehat{\boldsymbol{m}}_h^n = [\widehat{m}_x^n, \widehat{m}_y^n, \widehat{m}_z^n]^\top$ are handled with scaled mass matrices $\widehat{M}_* \in \mathbb{R}^{N\times N}$, where $\widehat{m}_{ij} = \langle \widehat{m}_*^n \phi_i, \phi_j\rangle$ for $* \in \{x, y, z\}$.

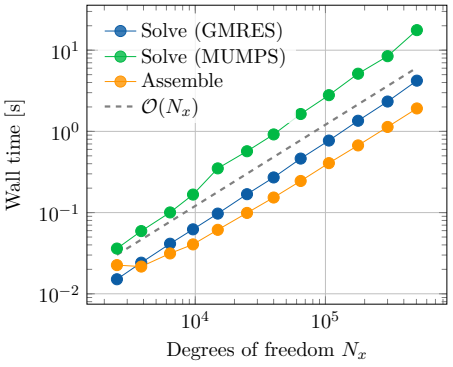


Figure 5.4.1: Total wall time per time-step for solving and assembling example (5.3.5) using MUMPS and GMRES with a SparseILU preconditioner and the previous solution as the initial guess. The iterative solver is terminated once the residual norm is below 10^{-6}.

We denote the nodal vectors of $\boldsymbol{v}_h^n$ and the Lagrange multiplier λ_h^n as $\underline{\boldsymbol{v}}^n \in \mathbb{R}^{3N}$ and $\underline{\lambda}^n \in \mathbb{R}^N$. The magnetization vector is represented by $\underline{\boldsymbol{m}}^n = [\underline{m}_x^n, \underline{m}_y^n, \underline{m}_z^n]^\top \in \mathbb{R}^{3N}$, where $\underline{m}_x^n, \underline{m}_y^n, \underline{m}_z^n \in \mathbb{R}^N$ corresponds to its spatial components. In total, this leads to the block matrix structure

$$\begin{bmatrix} \alpha M + \ell_{\mathrm{ex}}^2 \frac{\tau_n}{\delta_0} S & -\widehat{M}_z & \widehat{M}_y & \widehat{M}_x \\ \widehat{M}_z & \alpha M + \ell_{\mathrm{ex}}^2 \frac{\tau_n}{\delta_0} S & -\widehat{M}_x & \widehat{M}_y \\ -\widehat{M}_y & \widehat{M}_x & \alpha M + \ell_{\mathrm{ex}}^2 \frac{\tau_n}{\delta_0} S & \widehat{M}_z \\ \widehat{M}_x & \widehat{M}_y & \widehat{M}_z & 0 \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{v}}^n \\ \underline{\lambda}^n \end{bmatrix} = \begin{bmatrix} M\underline{f}_x^n + \sum_{j=1}^k \frac{\delta_j}{\delta_0} S\underline{m}_x^{n-j} \\ M\underline{f}_y^n + \sum_{j=1}^k \frac{\delta_j}{\delta_0} S\underline{m}_y^{n-j} \\ M\underline{f}_z^n + \sum_{j=1}^k \frac{\delta_j}{\delta_0} S\underline{m}_z^{n-j} \\ \underline{0} \end{bmatrix}.$$

We summarize the computational performance in Figure 5.4.1 by comparing the wall time required for assembling and solving with the MUMPS direct solver and the GMRES iterative solver with a SparseILU preconditioner, cf. Section 5.1. Since the scaled mass matrices depend on previous solutions, we reassemble the system at every time-step. Furthermore, both the assembling and solving times exhibit approximately linear growth with respect to the degrees of freedom. The average number of iterations required to reach a residual tolerance of 10^{-6} is provided in Figure 5.4.2.

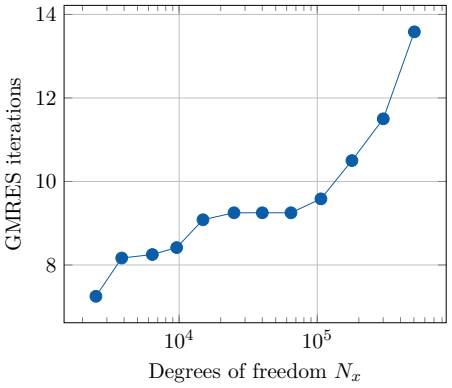


Figure 5.4.2: Average GMRES iterations per solve for example (5.3.5), using a SparseILU preconditioner and the previous time-step solution as the initial guess. The iterative solver is terminated once the residual norm is below 10^{-6}.

5.5 MOVING BUMP

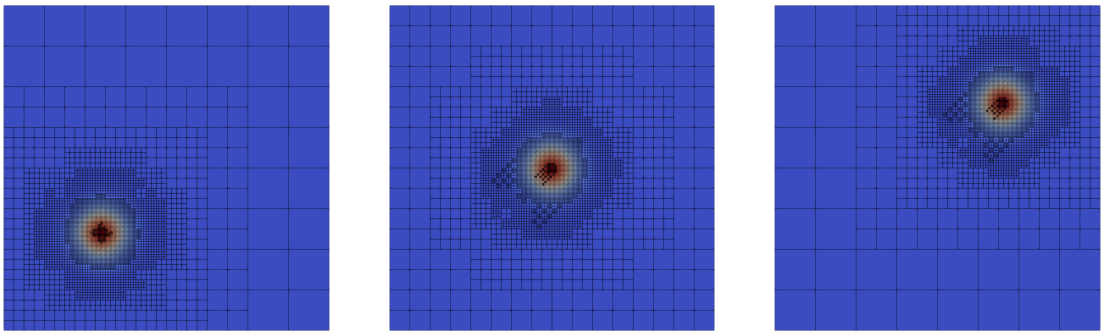
In this section, we study the example from [23, Sct. 3.7], which describes a localized bump propagating through the magnetic domain $\Omega = (0,1) \times (0,1)$. We choose the damping parameter $\alpha = 1$ and the exchange length $\ell_{\mathrm{ex}} = 1$. We set the final simulation time $T = 0.4$ and define the constant $C = 200$. Further, we apply the space and time adaptive algorithm described in Section 4.6 with BDF(1-4) and polynomial degree $p = 3$. We employ the face residual error estimator (4.6.1) and set the tolerances $\mathrm{TOL_t} = 10^{-7}$ and $\mathrm{TOL_s} = 10^{-4}$. We prescribe the exact solution $\boldsymbol{m} : [0,T] \times \Omega \to \mathbb{R}^3$ by

$$\boldsymbol{m}(t,\boldsymbol{x}) \coloneqq \frac{\sqrt{2}}{2} \begin{bmatrix} 0 \\ \sqrt{2 - \mathrm{e}^{-2Cr(t,\boldsymbol{x})}} \\ \mathrm{e}^{-Cr(t,\boldsymbol{x})} \end{bmatrix}, \tag{5.5.1}$$

for $(t,\boldsymbol{x}) \in [0,T] \times \Omega$, where

$$r(t,\boldsymbol{x}) \coloneqq (x_1 - t - t_0)^2 + (x_2 - t - t_0)^2$$

for $(t,\boldsymbol{x}) \in [0,T] \times \Omega$, where $t_0 = 0.3$. Obviously, we have $|\boldsymbol{m}(t,\boldsymbol{x})| = 1$ for $(t,\boldsymbol{x}) \in [0,T] \times \Omega$. Again, we compute the right-hand side $\boldsymbol{f}$ such that $\boldsymbol{m}$ defined in (5.3.5) is the exact solution



(a) $t = 0$ (b) $t = T/2$ (c) $t = T$

Figure 5.5.1: Evolution of the spatial mesh for example (5.5.1) with the space and time adaptive Algorithm 59. The background color represents the z-component of the magnetization.

of the alternative form (1.5.1) of the LLG equation, that is,

$$\boldsymbol{f} = \alpha \partial_t \boldsymbol{m} + \boldsymbol{m} \times \partial_t \boldsymbol{m} - \ell_{\mathrm{ex}}^2 \Delta \boldsymbol{m},$$

where we can neglect the orthogonal projection onto the tangent plane since we are solving in the tangent space anyway.

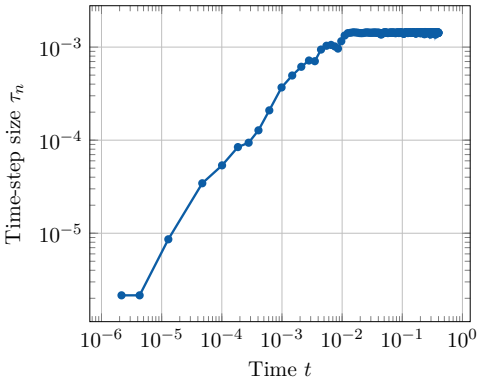
In addition to the $\boldsymbol{\mathcal{H}}^1$-error (5.3.1), we consider the normalization error

$$\| |\boldsymbol{m}_h^n| - 1 \|_{\mathcal{L}^\infty}, \tag{5.5.2}$$

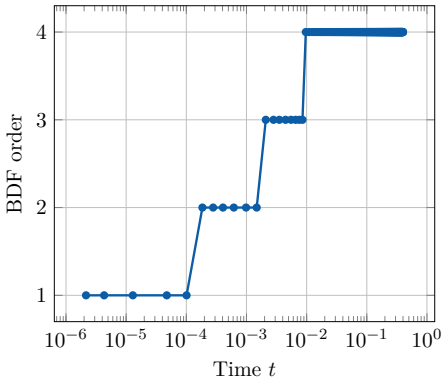
since we do not employ an explicit normalization step for $\boldsymbol{m}_h^n$. As a representative time a posteriori error estimator, we choose $\varepsilon_{2,n}$ given in (4.1.16), which is closely related to the global error indicator $\widetilde{\mathcal{E}}_2 = \sum_{n=2}^{N} \tau_n \varepsilon_{2,n}$ from Definition 28. As illustrated in Figure 5.5.1, the adaptive algorithm reliably locates the position of the bump and adjusts the spatial mesh accordingly. Throughout the simulation, the adaptive solver maintains a symmetric grid and preserves the symmetry of the solution.

Further details of the adaptive solver's performance are provided in Figure 5.5.2. Figure 5.5.2 (b) shows that the BDF order increases steadily until it stabilizes at order 4. Simultaneously, the time-step size does not exhibit any oscillations, see Figure 5.5.2 (a). Instead, it transitions smoothly and remains essentially constant once the initial adaptive phase is complete. We observe in Figure 5.5.2 (c), that the time-step suggestions for each order $k = 1, \dots, 4$ remain stable, while the spatial degrees of freedom stay approximately at $2.1 \cdot 10^5$, see Figure 5.5.2 (d). The a posteriori error estimators, the $\boldsymbol{\mathcal{H}}^1$-error and the normalization error all stay nearly constant over time, see Figure 5.5.3. Moreover, in Table 5.5.1, we summarize the total computation times for the space and time adaptive algorithm. The results demonstrate that the computational cost of the adaptive components, including error estimation and marking, is negligible compared to the linear solver time, which accounts for the majority of the overall runtime. This confirms that the adaptive strategy introduces only a

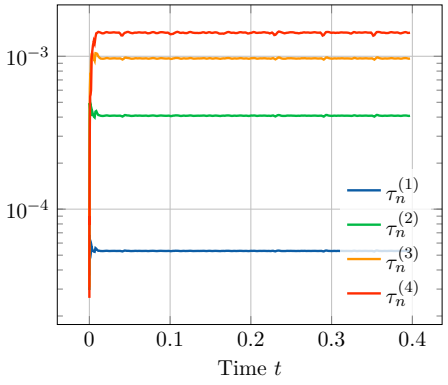
modest overhead while providing the benefits of adjusting the spatial mesh and the time-step size dynamically.



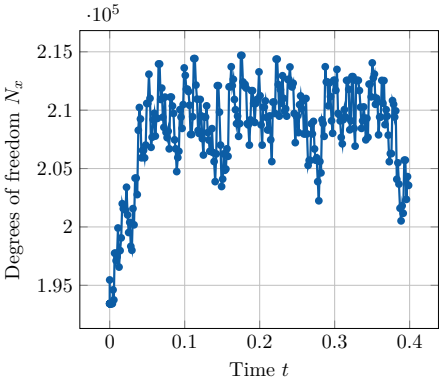
(a) Evolution of the time-step size



(b) Evolution of the BDF order



(c) Estimated time-step sizes



(d) Evolution of the degrees of freedom

Figure 5.5.2: Evolution of the time-step size (top left), the selected BDF order (top right), the estimated time-step size for BDF(1 - 4) (bottom left) and the spatial degrees of freedom (bottom right) for the moving bump example (5.5.1) using Algorithm 59.

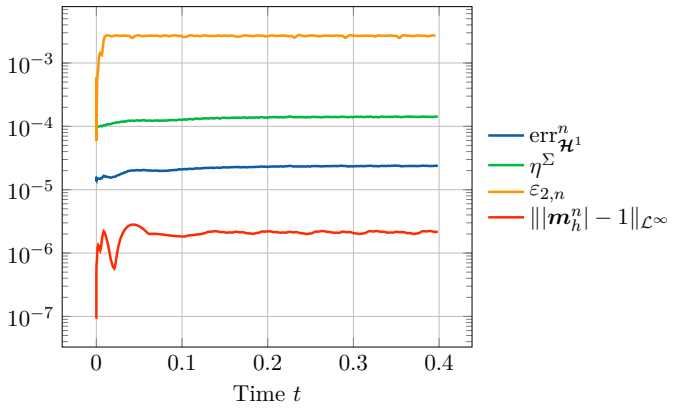


Figure 5.5.3: Evolution of the $\boldsymbol{\mathcal{H}}^1$-error (5.3.1), the global face residual error estimator η^{Σ} (4.6.3), the time a posteriori estimator $\varepsilon_{2,n}$ (4.1.16) and the normalization error $\||\boldsymbol{m}^n_h| - 1\|_{\mathcal{L}^\infty}$ (5.5.2) for the moving bump example (5.5.1) using Algorithm 59.

Solver Component	Time per Step (s)	Total Wall Time (s)
Assembly	0.70	226
Linear Solver	6.49	1910
Marking	0.00	0.16
Spatial Estimator	0.02	4.36
Time Estimator	0.01	3.12
Total	**7.28**	**2143.64**

Table 5.5.1: The computational cost per solver component for the moving bump example (5.5.1) with $\mathrm{TOL_t} = 10^{-7}$ and $\mathrm{TOL_s} = 10^{-4}$. The simulation involves a total of 61 064 300 degrees of freedom, with approximately 207 702 degrees of freedom per time-step. Spatial error estimation is performed using the face residual error estimator (4.6.1), the marking strategy as described in Section 4.4 and the time estimator consists of the VSVO BDF(1 - 4) algorithm introduced in Section 4.5.

5.5.1 Comparison between spatial error estimators

Additionally, we compare the global spatial error estimators defined in (4.6.3). Namely, we compare the gradient recovery estimator $\eta^{\boldsymbol{G}}(\boldsymbol{m}_h^n)$ (4.3.1) to the cell residual estimator $\eta^{\mathcal{T}}(\boldsymbol{m}_h^n)$ (4.6.2) and the face residual estimator $\eta^{\Sigma}(\boldsymbol{m}_h^n)$ (4.6.1). We compute the results with the final time $T = 0.01$ and the uniform BDF(2) method with the time-step size $\tau = T/100$.

		$\|\boldsymbol{\nabla}(\boldsymbol{m} - \boldsymbol{m}_h)\|_{\mathcal{L}^\infty(\boldsymbol{\mathcal{L}}^2)}$		$\max_n \eta^{\boldsymbol{G}}(\boldsymbol{m}_h^n)$		$\max_n \eta^{\mathcal{T}}(\boldsymbol{m}_h^n)$		$\max_n \eta^{\Sigma}(\boldsymbol{m}_h^n)$	
p	h	Value	Rate	Value	Rate	Value	Rate	Value	Rate
1	1/42	2.540e-01	—	2.418e-01	—	1.759e+00	—	1.116e+00	—
	1/54	1.958e-01	1.057	1.896e-01	0.989	1.394e+00	0.946	8.948e-01	0.897
	1/70	1.499e-01	1.046	1.469e-01	0.998	1.088e+00	0.971	7.038e-01	0.940
	1/91	1.148e-01	1.033	1.133e-01	1.002	8.427e-01	0.985	5.478e-01	0.967
	1/118	8.821e-02	1.022	8.755e-02	1.003	6.526e-01	0.993	4.255e-01	0.982
2	1/42	2.780e-02	—	1.302e-02	—	3.065e-01	—	7.487e-02	—
	1/54	1.682e-02	2.021	6.724e-03	2.656	1.857e-01	2.014	3.701e-02	2.833
	1/70	1.004e-02	2.005	3.354e-03	2.702	1.107e-01	2.010	1.783e-02	2.837
	1/91	5.946e-03	2.007	1.625e-03	2.779	6.548e-02	2.013	8.413e-03	2.881
	1/118	3.538e-03	2.008	7.790e-04	2.843	3.890e-02	2.014	3.958e-03	2.917
3	1/42	2.663e-03	—	2.085e-03	—	5.281e-02	—	1.707e-02	—
	1/54	1.310e-03	2.842	1.049e-03	2.750	2.604e-02	2.833	8.822e-03	2.644
	1/70	6.059e-04	2.987	4.983e-04	2.886	1.213e-02	2.960	4.280e-03	2.803
	1/91	2.768e-04	2.999	2.307e-04	2.948	5.558e-03	2.987	2.012e-03	2.889
	1/118	1.272e-04	3.002	1.067e-04	2.976	2.558e-03	2.997	9.400e-04	2.938

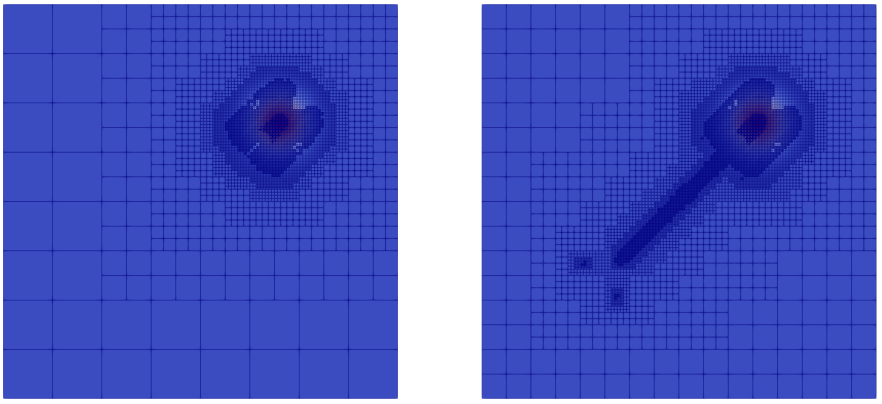
Table 5.5.2: Estimated errors of the gradient recovery estimator $\eta^{\boldsymbol{G}}(\boldsymbol{m}_h^n)$, the cell residual estimator $\eta^{\mathcal{T}}(\boldsymbol{m}_h^n)$ and the face residual estimator $\eta^{\Sigma}(\boldsymbol{m}_h^n)$ in comparison to the $\boldsymbol{\mathcal{H}}^1$-error for the moving bump example (5.5.1).

In Table 5.5.2, we provide a detailed comparison of these spatial error estimators for the polynomial degrees $p = 1, 2, 3$. Notably, the convergence rate for the gradient recovery estimator $\eta^{\boldsymbol{G}}(\boldsymbol{m}_h^n)$ behaves similarly to the face residual estimator $\eta^{\Sigma}(\boldsymbol{m}_h^n)$, yielding an additional order of convergence for the even polynomial degree $p = 2$ compared to the $\boldsymbol{\mathcal{H}}^1$-error. Although the gradient recovery estimator $\eta^{\boldsymbol{G}}(\boldsymbol{m}_h^n)$ is relatively close to the $\mathcal{H}^1$-error, it tends to slightly underestimate the actual values. Conversely, the cell residual estimator

$\eta^{\mathcal{T}}(\boldsymbol{m}_h^n)$ converges at the same rate as the $\boldsymbol{\mathcal{H}}^1$-error for all polynomial degrees.

5.5.2 NORMALIZATION AS POST-PROCESSING

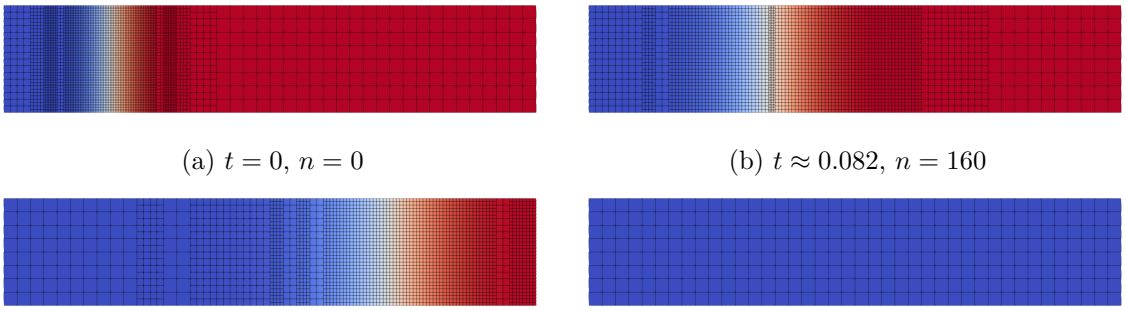
In this section, we investigate the evolution of the spatial mesh by comparing the impact of employing explicit normalization as a post-processing step before evaluating the spatial error estimator to refine and coarsen the mesh, compared to not applying the normalization. We set the final time $T = 0.4$ and take the polynomial degree $p = 1$. We apply the space and time adaptive Algorithm 59 using the face residual error estimator (4.6.1). It is important to emphasize that the normalized magnetization is utilized exclusively during post-processing for the spatial error estimator and does not influence the subsequent time-stepping algorithm.



(a) With normalization (b) Without normalization

Figure 5.5.4: Spatial mesh at final time $t = T = 0.4$ for the moving bump example (5.5.1). On the left, the grid is obtained by explicitly normalizing $\boldsymbol{m}_h^n$ as a post-processing step prior to evaluating the face residual error estimator (4.6.1). On the right, the mesh is generated without any post-processing. The background color represents the z-component of the magnetization.

In Figure 5.5.4, we illustrate the resulting spatial meshes at the final time. While the mesh with normalization coarsens effectively once the bump has passed, the version without normalization remains fine in the bump's wake and fails to coarsen. Because the mesh density around the active bump remains comparable in both cases, the primary advantage of post-processing normalization is the significantly improved coarsening behavior.



(a) $t = 0$, $n = 0$ (b) $t \approx 0.082$, $n = 160$

(c) $t \approx 0.23$, $n = 320$ (d) $t = T = 1.0$, $n = 590$

Figure 5.6.1: Evolution of the spatial mesh after n time-steps for example (5.6.1) with the space and time adaptive Algorithm 59. The background color represents the z-component of the magnetization, ranging from -1 to 1 (blue to red).

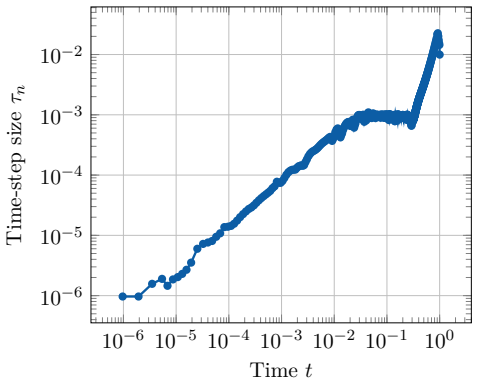
5.6 Domain wall

In this section, we simulate the motion of a domain wall. We consider the domain $\Omega = (0,1) \times (0,0.2)$ and fix the final simulation time $T = 1.0$. We fix $\ell_{\mathrm{ex}} = 0.1$ and $\alpha = 1$ and provide the initial data

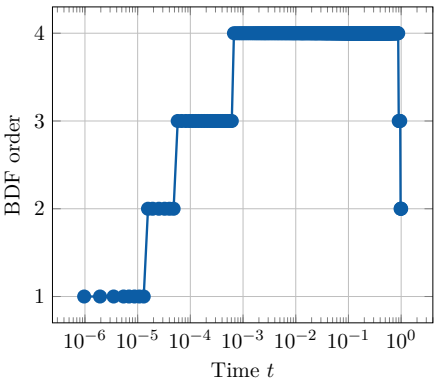
$$\boldsymbol{m}_0(\boldsymbol{x}) = \begin{cases} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} & \text{if } x_1 < c - r, \\ \begin{bmatrix} 0 \\ \cos(\pi\zeta(x_1)/2) \\ \sin(\pi\zeta(x_1)/2) \end{bmatrix} & \text{if } c - r \leq x_1 \leq c + r, \\ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & \text{if } c + r \leq x_1, \end{cases} \tag{5.6.1}$$

where $\zeta(x) \coloneqq \sin(\pi(x - c)/2r)$, $c = 0.2$ denotes the x-position of the interface and $r = 0.125$ its radius. An exact solution for this initial data is not known. Furthermore, we apply a constant external field $\boldsymbol{f} = [0, 0, -50]$. We apply the space and time adaptive algorithm described in Section 4.6 with BDF(1-4) and polynomial degree $p = 3$. We employ the face residual error estimator (4.6.1) and set the tolerances $\mathrm{TOL_t} = 10^{-7}$ and $\mathrm{TOL_s} = 10^{-4}$.

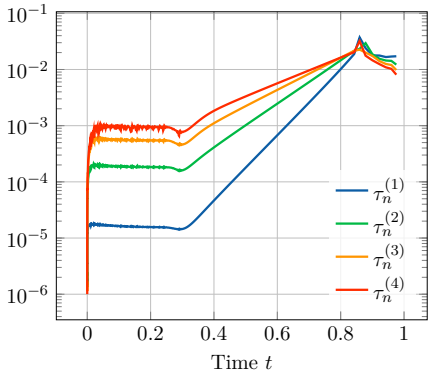
As illustrated in Figure 5.6.1, the adaptive algorithm ensures that the spatial mesh moves along with the domain wall as it progresses through the domain. At the end, the magnetization reaches a constant state with $\boldsymbol{m} = [0, 0, -1]^\top$ and the spatial mesh is consequently coarsened back to the underlying uniform mesh. Figure 5.6.2 provides further details on the adaptive solver's performance, showing in Figure 5.6.2 (b) that the BDF order increases steadily until it stabilizes at order 4. At approximately $t = 0.4$, the magnetization reaches a steady state, which results in a continuous increase in the time-step size (Figure 5.6.2 (a) and Figure 5.6.2 (c)) and significant mesh coarsening (Figure 5.6.2 (d)). Finally, Figure 5.6.3 presents the evolution of the a posteriori error estimators alongside the $\boldsymbol{\mathcal{H}}^1$-error and the normalization error, highlighting the consistent decay of the temporal estimator as the solution stabilizes.



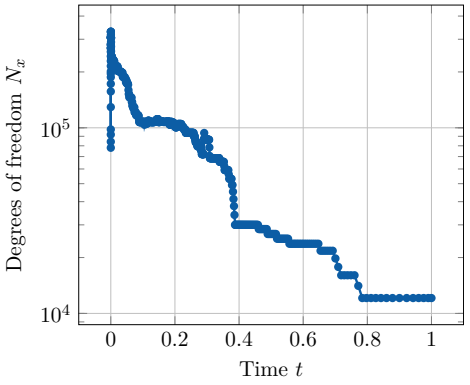
(a) Evolution of the time-step size



(b) Evolution of the BDF order



(c) Estimated time-step sizes



(d) Evolution of the degrees of freedom

Figure 5.6.2: Evolution of the time-step size (top left), the selected BDF order (top right), the estimated time-step size for BDF(1 - 4) (bottom left) and the spatial degrees of freedom (bottom right) for the domain wall example (5.6.1) using Algorithm 59.

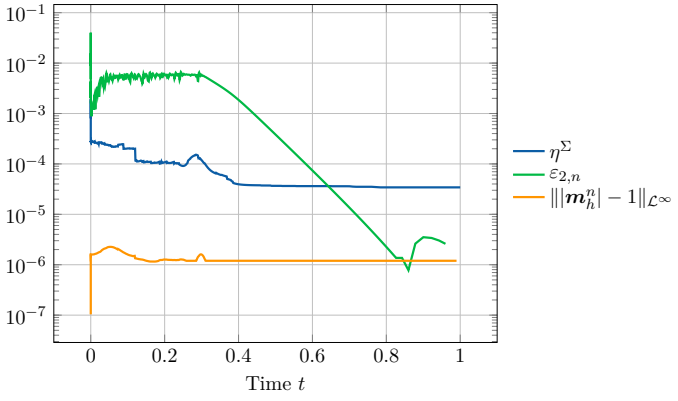


Figure 5.6.3: Evolution of the global face residual error estimator η^{Σ} (4.6.3), the time a posteriori estimator $\varepsilon_{2,n}$ (4.1.16) and the normalization error $\||\boldsymbol{m}_h^n| - 1\|_{\mathcal{L}^\infty}$ (5.5.2) for the domain wall example (5.6.1) using Algorithm 59.

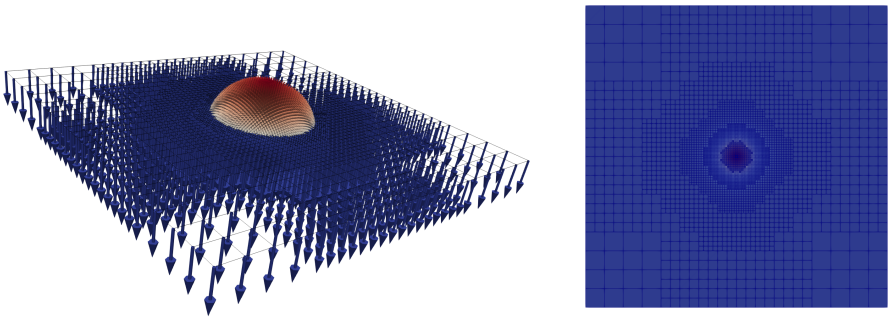


Figure 5.7.1: Initial solution (left) and the corresponding mesh (right) for the example (5.7.1). The color represents the z-component of the magnetization, ranging from -1 to 1 (blue to red).

5.7 A SINGULAR EXAMPLE?

Here, we consider an example that appears to show singular behavior and has been discussed in [22, 28, 29, 44]. In [36], we showed that computations on fine meshes and particularly adaptivity prevent a singularity from forming. We let $\Omega = (-1/2, 1/2)^2$ and choose constants $\ell_{\text{ex}} = 1$ and $\alpha = 1$. We apply the space and time adaptive algorithm described in Section 4.6 with BDF(1-4) using polynomial degrees $p = 1$ or $p = 3$ in space. We set the final time $T = 2.0$. We employ the stabilized order control described in Section 4.5 due to the lack of regularity expected in the solution. We utilize the face residual error estimator (4.6.1), fix the time tolerances $\text{TOL}_\text{t} = 10^{-7}$ and apply various TOL_s specified later. We prescribe the initial data

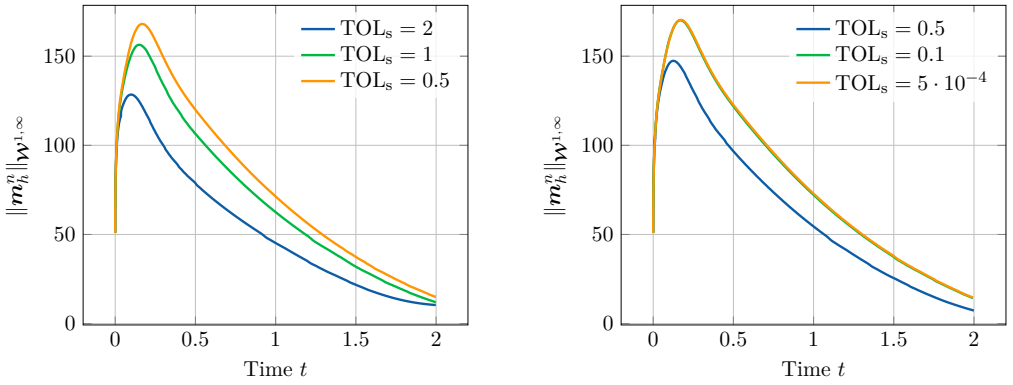
$$\boldsymbol{m}_0(\boldsymbol{x}) = \begin{cases} \frac{1}{A(\boldsymbol{x})^2+|\boldsymbol{x}|^2} \begin{bmatrix} 2A(\boldsymbol{x})\boldsymbol{x} \\ A(\boldsymbol{x})^2-|\boldsymbol{x}|^2 \end{bmatrix} & \text{if } |\boldsymbol{x}|^2 \leq 1/2\,, \\ \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} & \text{otherwise,} \end{cases} \tag{5.7.1}$$

where $A(\boldsymbol{x}) = (1 - 2|\boldsymbol{x}|)^4/s$ and $s = 16$. Simple computations show that $|\boldsymbol{m}_0(\boldsymbol{x})| = 1$ for $\boldsymbol{x} \in \Omega$. We apply no external field, i.e., $\boldsymbol{f} = \boldsymbol{0}$. Since an analytical solution is not available, we must rely on the qualitative behavior of the numerical solution. For this example, we prohibit coarsening and set $\theta_\text{c} = 0$.

The evolution of the $\boldsymbol{\mathcal{W}}^{1,\infty}$-norm of the numerical solution is illustrated in Figure 5.7.2. It is important to note that no blow-up occurs in any of the simulations. Instead, the solution saturates at a maximum gradient of approximately 170. More specifically, we observe that $\max_{0\leq n\leq N} \|\boldsymbol{m}_h^n\|_{\boldsymbol{\mathcal{W}}^{1,\infty}} \approx 170$.

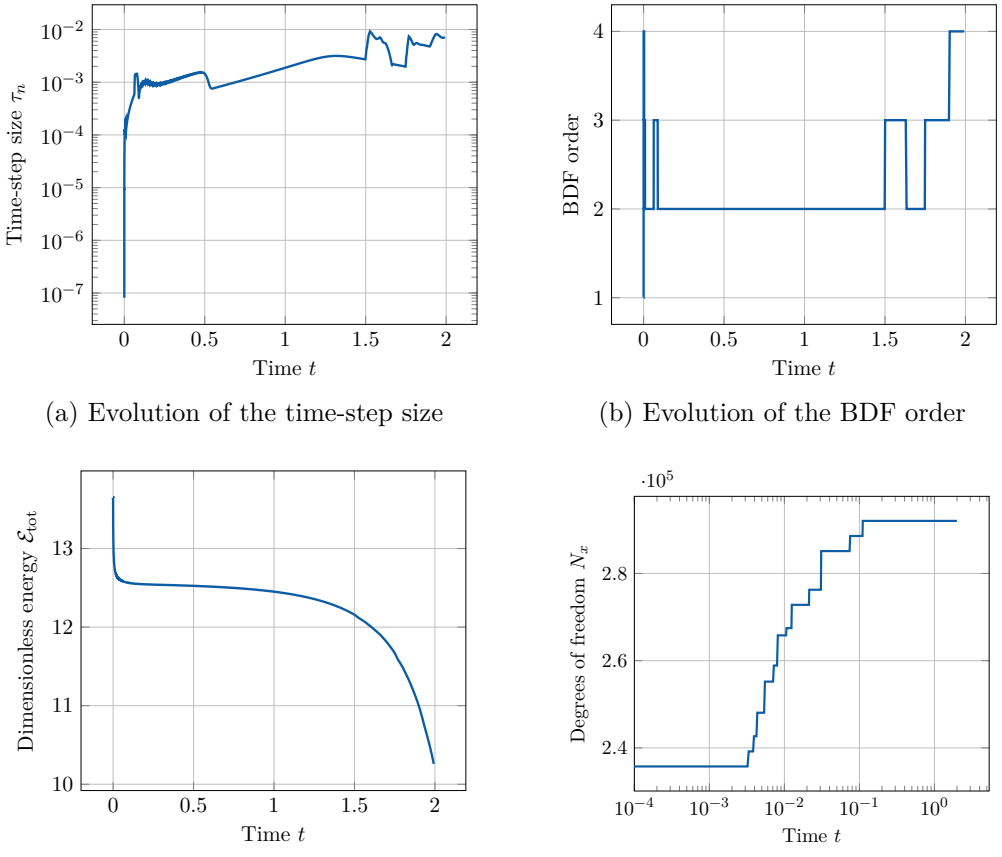
Figure 5.7.3 provides more details of the adaptive algorithm for the specific case of the degree of polynomial $p = 3$ in space with $\text{TOL}_\text{s} = 5\cdot 10^{-4}$. The total (dimensionless) energy, consisting only of the exchange contribution $\mathcal{E}_\text{tot}(t) = \mathcal{E}_\text{ex}(t) = \frac{1}{2}\ell_\text{ex}^2\|\boldsymbol{\nabla m}(t)\|^2$, decays monotonically, as illustrated in Figure 5.7.3 (c). Up to $t \approx 1.5$, the VSVO algorithm maintains the order 2 (Figure 5.7.3 (b)) while producing a smooth sequence of time-step sizes (Figure 5.7.3 (a)). During the initial phase up to $t \approx 0.1$, the spatial mesh is continuously refined from $\approx 240\,000$

to ≈ 290 000 degrees of freedom, see Figure 5.7.3 (d). Toward the end of the simulation, the algorithm transitions to the fourth-order BDF scheme (Figure 5.7.3 (b)).



(a) $p = 1$ (b) $p = 3$

Figure 5.7.2: Evolution of $\|m_h^n\|_{W^{1,\infty}}$ for the example (5.7.1) using Algorithm 59.



(a) Evolution of the time-step size (b) Evolution of the BDF order

(c) Evolution of the total energy (d) Evolution of the degrees of freedom

Figure 5.7.3: Evolution of the time-step size (top left), the selected BDF order (top right), the total energy (bottom left) and the spatial degrees of freedom (bottom right) for the example (5.7.1) using Algorithm 59 with cubic polynomials and TOL$_s$ = $5 \cdot 10^{-4}$.

COMPARISON TO THE POINTWISE DISCRETE TANGENT SPACE FORMULATION

We also consider example (5.7.1) for a tangent plane scheme that enforces the normalization constraint in a pointwise sense. As illustrated in the following, this formulation exhibits a blow-up for larger tolerances, whereas the blow-up is avoided when the tolerances are sufficiently small. Importantly, the numerical solution of both formulations eventually saturate toward the same solution, which does not blow up.

Let $V_h^n = \mathcal{S}_h^1(\mathcal{T}_h^n)$ be the space of globally continuous piecewise linear polynomials on the triangulation $\mathcal{T}_h^n$. Then, define

$$\boldsymbol{T}_{h,\mathrm{pw}}^n(\boldsymbol{m}_h^n) \coloneqq \{\boldsymbol{\varphi}_h^n \in \boldsymbol{V}_h^n \ : \ \boldsymbol{m}_h^n(\boldsymbol{z}) \cdot \boldsymbol{\varphi}_h^n(\boldsymbol{z}) = 0 \text{ for all } \boldsymbol{z} \in \mathcal{N}_h^n\}, \tag{5.7.2}$$

where $\mathcal{N}_h^n$ denotes the set of all nodes in $\mathcal{T}_h^n$. This specific type of discrete tangent space has been employed, for instance, in [11–13, 16, 30, 107]. Then, the fully discrete scheme seeks $\boldsymbol{m}_h^n \in \boldsymbol{V}_h^n$ with $\boldsymbol{v}_h^n \in \boldsymbol{T}_{h,\mathrm{pw}}^n(\widehat{\boldsymbol{m}}_h^n)$ such that

$$\alpha\langle \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \langle \widehat{\boldsymbol{m}}_h^n \times \boldsymbol{v}_h^n, \boldsymbol{\varphi}_h^n\rangle + \ell_{\mathrm{ex}}^2 \langle \boldsymbol{\nabla} \boldsymbol{m}_h^n, \boldsymbol{\nabla} \boldsymbol{\varphi}_h^n\rangle = \langle \boldsymbol{f}_h^n, \boldsymbol{\varphi}_h^n\rangle, \tag{5.7.3}$$

for all $\boldsymbol{\varphi}_h^n \in \boldsymbol{T}_{h,\mathrm{pw}}^n(\widehat{\boldsymbol{m}}_h^n)$. Unlike (1.6.6), the normalization constraint is now enforced pointwise at each node rather than in the $\mathcal{L}^2$-sense. For details of the implementation of the discrete pointwise tangent space, we refer to [107, Sct. 6.1].

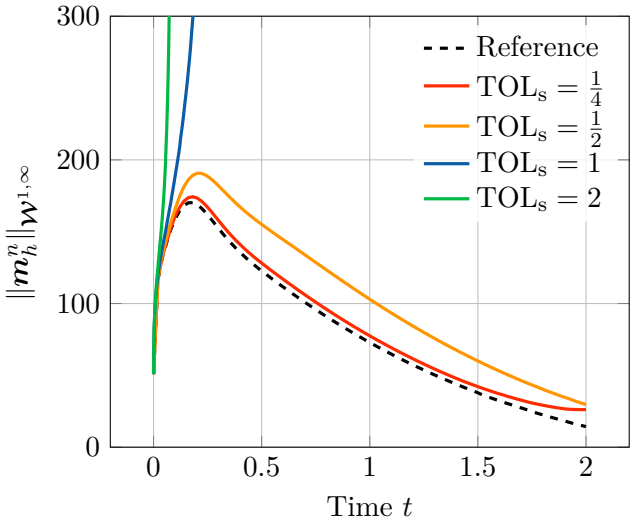
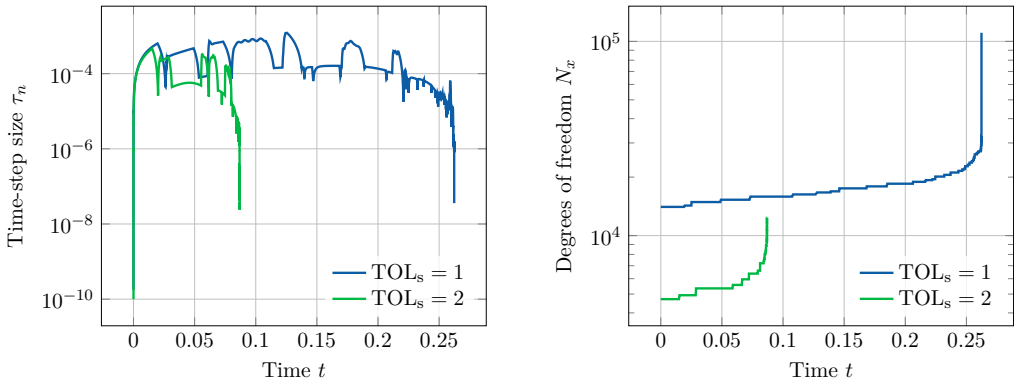


Figure 5.7.4: Evolution of the $\boldsymbol{\mathcal{W}}^{1,\infty}$-norm of the discrete solution for the example (5.7.1) using Algorithm 59 with the pointwise discrete tangent space formulation (5.7.3). The reference solution was computed with cubic polynomials and $\mathrm{TOL_s} = 5 \cdot 10^{-4}$ as detailed in Figure 5.7.3.

For (5.7.3) we restrict the simulations to linear finite elements as higher-order convergence is generally not observed for this specific discrete tangent space [15, 16]. Although this can be corrected by adding a correction term as suggested in [7, Remark 2.2], here we focus on linear finite elements. The temporal discretization is again done by the BDF scheme and we apply

the same space and time adaptive Algorithm 59 with BDF(1 - 4). We fix the time tolerance at $\mathrm{TOL_t} = 10^{-7}$ and apply the space tolerances $\mathrm{TOL_s} \in \{2, 1, \frac{1}{2}, \frac{1}{4}\}$ using the face residual estimator (4.6.1).



(a) Evolution of the time-step size (b) Evolution of the degrees of freedom

Figure 5.7.5: Evolution of the time-step size (left) and the spatial degrees of freedom (right) for the example (5.7.1) using Algorithm 59 with the pointwise discrete tangent space formulation (5.7.3).

Figure 5.7.4 shows that the numerical solution blows up for $\mathrm{TOL_s} = 2$ and $\mathrm{TOL_s} = 1$ in the sense that the $\boldsymbol{\mathcal{W}}^{1,\infty}$-norm of the numerical solution tends to infinity. In contrast, in Figure 5.7.2 (a) we do not observe a blow-up at these tolerances. The adaptive algorithm attempts to mitigate this blow-up by drastically reducing the time-step size and increasing the spatial degrees of freedom, as illustrated in Figure 5.7.5. However, for smaller tolerances $\mathrm{TOL_s} = \frac{1}{2}$ and $\mathrm{TOL_s} = \frac{1}{4}$, the adaptive algorithm successfully avoids blow-up. The numerical solution of (5.7.3) eventually saturates toward the reference solution shown in Figure 5.7.3, which we obtained with cubic polynomials and the fine spatial tolerance $\mathrm{TOL_s} = 5 \cdot 10^{-4}$. In the end, the numerical solution does not blow up toward $[0, 0, -1]$, but instead moves slowly toward $[0, 0, 1]$.

5.8 FINITE ELEMENT APPROXIMATION OF THE STRAY FIELD

In this section, we describe the computation of the stray field using boundary element methods. Following the comprehensive treatment of boundary integral equations in [110], we briefly recall the theory of elliptic boundary integral operators. Finally, we derive a finite element approximation for the transmission problem (1.2.2). This section is based on [76, 105, 110].

Let $\Omega \subset \mathbb{R}^3$ be a bounded Lipschitz domain and let G denote the fundamental solution of the Laplace operator in three dimensions, given by

$$G(\boldsymbol{x} - \boldsymbol{y}) = -\frac{1}{4\pi} \frac{1}{|\boldsymbol{x} - \boldsymbol{y}|}, \qquad \boldsymbol{\nabla}_{\boldsymbol{y}} G(\boldsymbol{x} - \boldsymbol{y}) = \frac{1}{4\pi} \frac{(\boldsymbol{x} - \boldsymbol{y})}{|\boldsymbol{x} - \boldsymbol{y}|^3}.$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$ with $\boldsymbol{x} \neq \boldsymbol{y}$. Then, for a function $v \in \mathcal{L}^1(\partial\Omega)$, the associated double layer

potential is defined by

$$(\mathrm{DL}\, v)(\boldsymbol{x}) \coloneqq \int_{\partial\Omega} \boldsymbol{\nabla}_{\boldsymbol{y}} G(\boldsymbol{x}-\boldsymbol{y}) \cdot \boldsymbol{n}(\boldsymbol{y}) v(\boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}) \tag{5.8.1}$$

for $\boldsymbol{x} \in \mathbb{R}^3 \setminus \partial\Omega$. Since the fundamental solution G is regular for $\boldsymbol{x} \neq \boldsymbol{y}$, the double layer potential is well defined for $\boldsymbol{x} \in \mathbb{R}^3 \setminus \partial\Omega$. Moreover, the double layer potential is harmonic, as stated in the following theorem, cf. [110, Thm. 3.1.1].

Theorem 62. *Let $v \in \mathcal{L}^1(\partial\Omega)$. Then, the double layer potential satisfies*

$$-\Delta(\mathrm{DL}\, v) = 0$$

in $\mathbb{R}^3 \setminus \partial\Omega$ in a distributional sense.

If we assume additional smoothness of the surface, then the double layer potential is well-defined on the boundary, since the integrand in (5.8.1) is improperly integrable [110, Lemma 3.3.8]. This is summarized in the following theorem.

Lemma 63. *Let Ω be a bounded Lipschitz domain with $\partial\Omega \in \mathcal{C}^2_{pw}$. Then, the mapping $D : \mathcal{L}^\infty(\partial\Omega) \to \mathcal{L}^\infty(\partial\Omega)$ with*

$$(Dv)(\boldsymbol{x}) \coloneqq \int_{\partial\Omega} \boldsymbol{\nabla}_{\boldsymbol{y}} G(\boldsymbol{x}-\boldsymbol{y}) \cdot \boldsymbol{n}(\boldsymbol{y}) v(\boldsymbol{y})\, \mathrm{d}s(\boldsymbol{y}), \qquad \boldsymbol{x} \in \partial\Omega,$$

is continuous.

The characteristic behavior of the double layer potential on the surface $\partial\Omega$ is described by the following jump relations, cf. [110, Thm 3.3.1].

Theorem 64 (Jump properties)**.** *Let Ω be a bounded Lipschitz domain. Then, the double layer potential satisfies the jump properties*

$$[\mathrm{DL}\, \psi] = \psi \text{ in } \mathcal{H}^{1/2}(\partial\Omega), \qquad [\boldsymbol{\nabla}\, \mathrm{DL}\, \psi \cdot \boldsymbol{n}] = 0 \text{ in } \mathcal{H}^{-1/2}(\partial\Omega) \tag{5.8.2}$$

for all $\psi \in \mathcal{H}^{1/2}(\partial\Omega)$.

For a domain with smooth boundary, the following jump properties hold, cf. [110, Thm 3.3.13].

Theorem 65 (Jump properties on a smooth boundary)**.** *Let Ω be a bounded Lipschitz domain with $\partial\Omega \in \mathcal{C}^2_{pw}$. Then, the traces of the double layer potential from the exterior and interior domains satisfy*

$$\mathrm{DL}\, v(\boldsymbol{x})|_{\pm} = \pm\frac{1}{2} v(\boldsymbol{x}) + (Dv)(\boldsymbol{x}), \qquad \boldsymbol{x} \in \partial\Omega, \tag{5.8.3}$$

where

$$v(\boldsymbol{x})|_{+} = \lim_{\substack{\boldsymbol{y}\to\boldsymbol{x} \\ \boldsymbol{y}\in\mathbb{R}^3\setminus\overline{\Omega}}} v(\boldsymbol{y}), \qquad v(\boldsymbol{x})|_{-} = \lim_{\substack{\boldsymbol{y}\to\boldsymbol{x} \\ \boldsymbol{y}\in\Omega}} v(\boldsymbol{y}).$$

FEM-BEM approach by Fredkin and Koehler

For convenience, we begin by recalling the transmission problem (1.2.2) in a dimensionless form. For a bounded domain $\Omega \subset \mathbb{R}^3$, we seek a potential $u \in \mathcal{H}^1(\mathbb{R}^3)$ that satisfies

$$\begin{aligned} -\Delta u^{\text{int}} &= -\boldsymbol{\nabla} \cdot \boldsymbol{m}, && \text{in } \Omega, \\ -\Delta u^{\text{ext}} &= 0, && \text{in } \mathbb{R}^3 \setminus \overline{\Omega}, \\ u^{\text{ext}} - u^{\text{int}} &= 0, && \text{on } \partial\Omega, \\ \boldsymbol{\nabla}(u^{\text{ext}} - u^{\text{int}}) \cdot \boldsymbol{n} &= -\boldsymbol{m} \cdot \boldsymbol{n}, && \text{on } \partial\Omega, \\ u(\boldsymbol{x}) &= \mathcal{O}(1/|\boldsymbol{x}|), && \text{for } |\boldsymbol{x}| \to \infty, \end{aligned} \tag{5.8.4}$$

where the superscripts 'ext' and 'int' refer to the traces of u on $\partial\Omega$ with respect to the exterior domain $\mathbb{R}^3 \setminus \overline{\Omega}$ and the interior domain Ω, respectively. Throughout this section, $\boldsymbol{n}$ denotes the unit outer normal vector relative to Ω. The (dimensionless) stray field is subsequently obtained by $\boldsymbol{h}_{\text{s}} = \boldsymbol{H}_{\text{s}}/M_s = -\boldsymbol{\nabla} U/M_s = -\boldsymbol{\nabla} u$.

We consider the approach proposed by Fredkin and Koehler [70, 71] to solve the transmission problem (5.8.4), which is particularly well-suited for finite element approximations. Moreover, we exploit the fact that we only need to compute the stray field for a bounded domain instead of the full space $\mathbb{R}^3$. The key idea is to decompose the transmission problem (5.8.4) into two separate problems, namely an inhomogeneous Poisson problem with Neumann boundary conditions and a homogeneous Poisson problem with Dirichlet boundary conditions. Specifically, we decompose the scalar potential by $u = u_1 + u_2$, where $u_1 \in \mathcal{H}^1(\Omega)$ is the solution of the Neumann problem

$$\begin{aligned} -\Delta u_1 &= -\boldsymbol{\nabla} \cdot \boldsymbol{m}, && \text{in } \Omega, \\ \boldsymbol{\nabla} u_1 \cdot \boldsymbol{n} &= \boldsymbol{m} \cdot \boldsymbol{n}, && \text{on } \partial\Omega. \end{aligned} \tag{5.8.5}$$

We can extend u_1 outside of $\overline{\Omega}$ by zero to define u_1 on $\mathbb{R}^3$. Next, we seek u_2 to ensure that $u_1 + u_2$ is a solution of the transmission problem (5.8.4). Thus, $u_2 = (u_2|_{\Omega}, u_2|_{\mathbb{R}^3\setminus\overline{\Omega}}) \in \mathcal{H}^1(\Omega) \times \mathcal{H}^1_{\text{loc}}(\mathbb{R}^3 \setminus \overline{\Omega})$ has to be the solution of

$$\begin{aligned} -\Delta u_2 &= 0, && \text{in } \mathbb{R}^3 \setminus \partial\Omega, \\ [u_2] &= u_1, && \text{on } \partial\Omega, \\ \boldsymbol{\nabla} u_2 \cdot \boldsymbol{n} &= 0, && \text{on } \partial\Omega, \\ u_2(\boldsymbol{x}) &= \mathcal{O}(1/|\boldsymbol{x}|), && \text{for } |\boldsymbol{x}| \to \infty. \end{aligned} \tag{5.8.6}$$

Here, the space $\mathcal{H}^1_{\text{loc}}(\mathbb{R}^3 \setminus \overline{\Omega})$ contains all continuous linear functionals in $\mathcal{C}_c^\infty(\mathbb{R}^3 \setminus \overline{\Omega})$, in short $v \in (\mathcal{C}_c^\infty(\mathbb{R}^3 \setminus \overline{\Omega}))^*$ with the property $\phi v \in \mathcal{H}^1(\mathbb{R}^3 \setminus \overline{\Omega})$ for all $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^3 \setminus \overline{\Omega})$, cf. [110, Def. 2.6.1].

Assuming $\Omega \subset \mathbb{R}^3$ with $\partial\Omega \in \mathcal{C}^2$, we obtain that $u_2 = \mathrm{DL}\, u_1$ solves (5.8.6) since the double layer potential is harmonic in $\mathbb{R}^3 \setminus \overline{\Omega}$, cf. Theorem 62, and satisfies the jump conditions (5.8.2).

Furthermore, using (5.8.3), the interior trace is given by

$$u_2(\boldsymbol{x})|_- = -\frac{1}{2}u_1(\boldsymbol{x})|_- + (Du_1)(\boldsymbol{x})$$

for $\boldsymbol{x} \in \partial\Omega$. Since we are only interested in u_2 in Ω, we obtain with (5.8.6) the interior Dirichlet problem

$$\begin{aligned} -\Delta u_2 &= 0, && \text{in } \Omega, \\ u_2 &= Du_1 - \frac{1}{2}u_1, && \text{on } \partial\Omega. \end{aligned} \tag{5.8.7}$$

If we only have $\partial\Omega \in \mathcal{C}^2_{\mathrm{pw}}$, we obtain instead with the functional

$$J(\boldsymbol{x}) \coloneqq -\int_{\partial\Omega} \boldsymbol{\nabla}_{\boldsymbol{y}} G(\boldsymbol{x}, \boldsymbol{y}) \cdot \boldsymbol{n}(\boldsymbol{y}) \,\mathrm{d}s(\boldsymbol{y}), \qquad \boldsymbol{x} \in \overline{\Omega}$$

the interior problem

$$\begin{aligned} -\Delta u_2 &= 0, && \text{in } \Omega, \\ u_2 &= Du_1 + (J-1)u_1, && \text{on } \partial\Omega, \end{aligned} \tag{5.8.8}$$

cf. [110, Lemma 3.3.11, Thm. 3.3.13].

We thus compute a finite element approximation of the potential u in the bounded domain Ω by first obtaining u_1 through a finite element approximation of the Neumann problem (5.8.5) and subsequently computing u_2 via a finite element approximation of the Dirichlet problem (5.8.8). To maintain numerical stability, we implement the boundary integrals in (5.8.8) using a Gaussian quadrature rule that avoids quadrature points directly on the degrees of freedom of the finite element space. Once the total potential $u = u_1 + u_2$ is determined in Ω, we obtain the stray field by $\boldsymbol{h}_{\mathrm{s}} = \boldsymbol{\nabla} u$. More details are provided in [76, Alg. 4.3.1] and [105, Alg. 12].

5.9 μMAG STANDARD PROBLEM #4

In this section, we compute the μMAG standard problem #4, which focuses on the dynamic aspects of micromagnetic computations, using the full adaptive algorithm described in Section 4.6. This benchmark is one of five benchmark problems defined by the Micromagnetic Modeling Activity Group [99], which is part of the Center for Theoretical and Computational Materials Science (CTCMS) at the National Institute of Standards and Technology (NIST), to address fundamental issues in micromagnetic modeling.

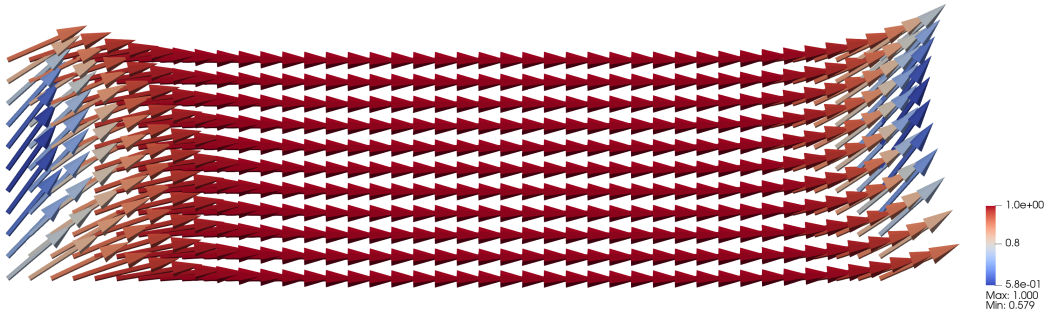


Figure 5.9.1: The x-y plane of the equilibrium s-state. The underlying color corresponds to the x-component of the magnetization.

The physical domain is a thin rectangular cuboid with a thickness of 3 nm, a length of 500 nm and a width of 125 nm. The exchange stiffness constant is set to $A = 1.3 \cdot 10^{-11}\,\mathrm{J/m}$ and the saturation magnetization is chosen as $M_s = 8 \cdot 10^5\,\mathrm{A/m}$. As reference parameters, we employ $L_{\mathrm{ref}} = 1\,\mathrm{nm}$ and $T_{\mathrm{ref}} = \gamma_0 M_s$, where $\gamma_0 \approx 2.21 \cdot 10^5\,\mathrm{m/(A\,s)}$ denotes the gyromagnetic ratio. The Gilbert damping parameter is given by $\alpha = 0.02$. We further recall that the permeability of vacuum is given by $\mu_0 = 4\pi \cdot 10^{-7}\,\mathrm{N/A^2}$. With these parameters, the exchange constant defined in (1.2.9) is evaluated to

$$\ell_{\mathrm{ex}}^2 = \frac{2A}{\mu_0 M_s^2 L_{\mathrm{ref}}^2} \approx 32.328. \tag{5.9.1}$$

Starting from an equilibrium s-state, see Figure 5.9.1, an external field of

$$\mu_0 \boldsymbol{H}_{\mathrm{ext}} = [-24.6, 4.3, 0]^\top\,\mathrm{mT} \tag{5.9.2}$$

is employed, which is sufficient to reverse the direction of the magnetization. In addition to the external field (5.9.2) and the exchange field with exchange constant (5.9.1), the stray field is included in the effective field. We compute the stray field with the FEM-BEM approach of Fredkin and Koehler [70, 71] as described in Section 5.8.

OBTAINING THE S-STATE

The initial state is an equilibrium s-state, see Figure 5.9.1. To obtain the s-state, we start with the uniform initial magnetization $\boldsymbol{m}_h^0 = [1, 0, 0]^\top$ and let $\boldsymbol{H}_{\mathrm{ext}}$ be an external field in the direction of $[1, 1, 1]^\top$ with a magnitude of $|\boldsymbol{H}_{\mathrm{ext}}|$ decreasing linearly from $30/\mu_0$ to 0 mT over

a period of 1 ns [103, 117]. To accelerate the process toward the equilibrium state, we choose $\alpha = 1$ for this period. After the external field $\boldsymbol{H}_{\text{ext}}$ decreased to zero after 1 ns, we relax the s-state for an additional period of 1 ns without applying any external field. Finally, we store the obtained s-state to serve as the initial data for the simulation.

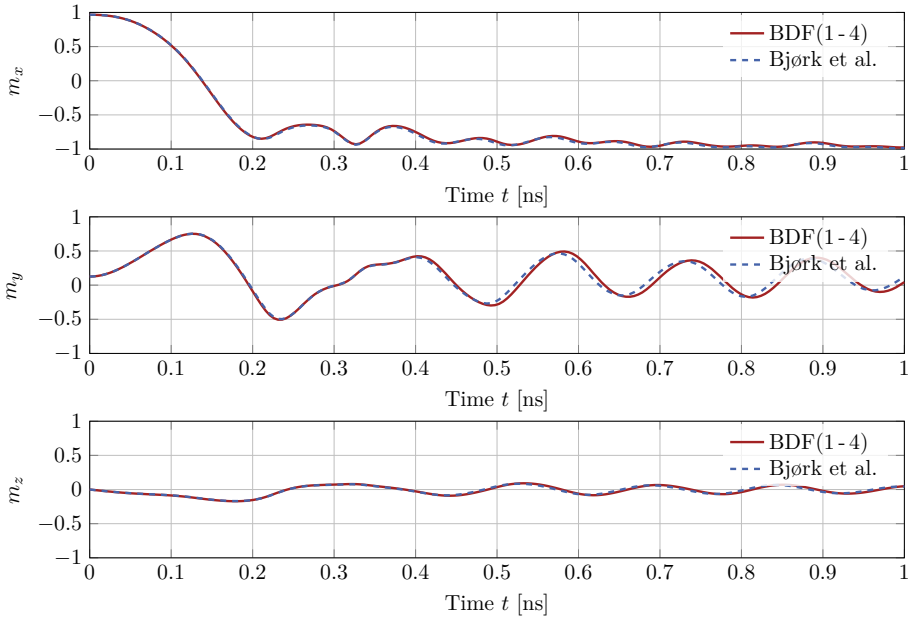
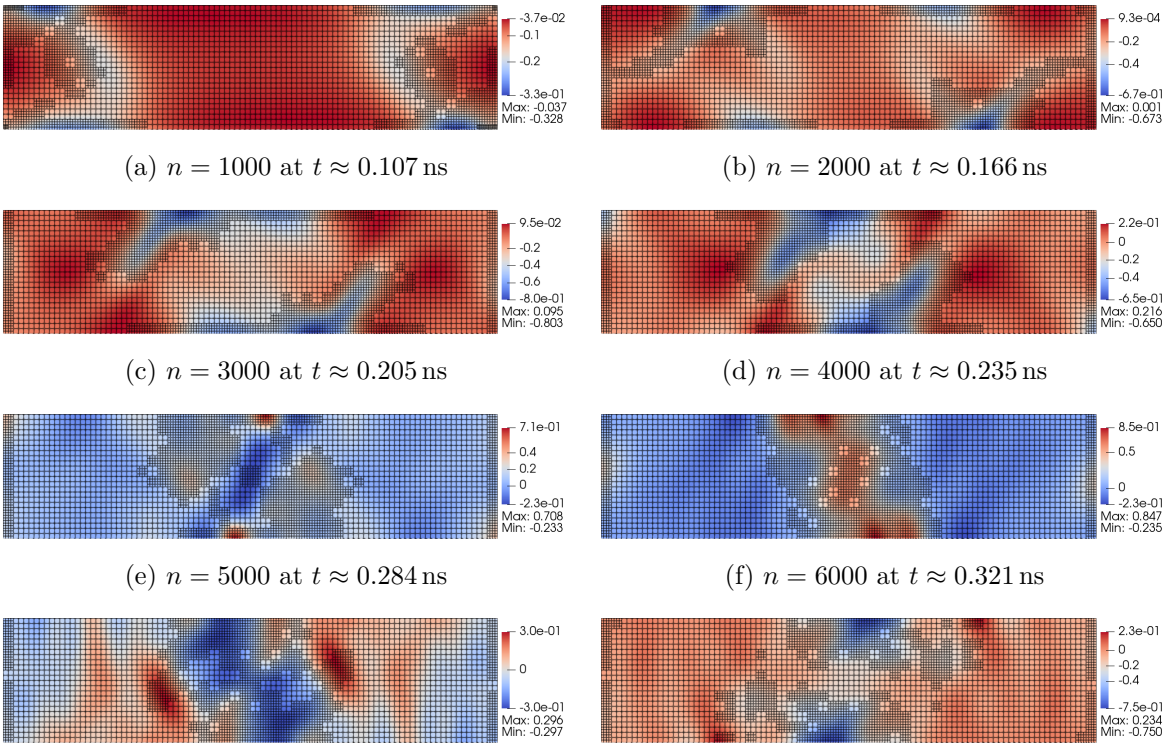


Figure 5.9.2: Evolution of the spatially averaged x-, y- and z-components of the magnetization for the space and time adaptive algorithm using the BDF(1 - 4) method with linear finite elements in comparison to a reference solution by Bjørk et al. for the μMAG standard problem #4.

ADAPTIVE SIMULATION

Starting from the equilibrium s-state, we compute the solution of the μMAG standard problem #4 using the full adaptive algorithm described in Section 4.6. To ensure numerical robustness in the presence of the small damping coefficient α, we employ the stabilized order selection introduced in Section 4.5. Time integration is performed with the VSVO BDF(1 - 4) algorithm. To reduce the spatial degrees in the z-direction, we employ linear finite elements and solve with the tangent plane scheme described in Section 1.6. As the final simulation time, we take $T = 1$ ns. We set the time tolerance $\text{TOL}_\text{t} = 10^{-3}$ and the spatial tolerance $\text{TOL}_\text{s} = 0.3$. As spatial error estimator, we employ the face residual error estimator (5.3.7). To limit computational complexity, the number of degrees of freedom is capped at $N_x^{\max} = 8 \cdot 10^4$. Consequently, spatial refinement is performed only if the degrees of freedom remain below $N_x^{\max}$. The adaptive refining process starts with an underlying uniform mesh consisting of $64 \times 16 \times 2$ cells.

We compare our results with the reference solution of Bjørk et al., which was computed with the open source micromagnetic simulation framework MagTense [35]. Figure 5.9.2 shows the evolution of the spatially averaged x-, y- and z-components of the magnetization in comparison



(a) $n = 1000$ at $t \approx 0.107\,\text{ns}$ (b) $n = 2000$ at $t \approx 0.166\,\text{ns}$

(c) $n = 3000$ at $t \approx 0.205\,\text{ns}$ (d) $n = 4000$ at $t \approx 0.235\,\text{ns}$

(e) $n = 5000$ at $t \approx 0.284\,\text{ns}$ (f) $n = 6000$ at $t \approx 0.321\,\text{ns}$

(g) $n = 7000$ at $t \approx 0.390\,\text{ns}$ (h) $n = 8000$ at $t \approx 0.436\,\text{ns}$

Figure 5.9.3: Illustration of the evolution of the spatial mesh after n time-steps for the space and time adaptive algorithm using the BDF(1 - 4) method with linear finite elements. The underlying color corresponds to the z-component of the magnetization.

to the reference solution. In principle, the absence of an explicit normalization step may induce a phase shift, as discussed in [76, Sct. 6.3] and [107, Sct. 6.2.2.1]. This effect is usually most evident in the spatially averaged y-component of the magnetization. Despite employing a projection-free scheme (except for the predictor described in Section 4.2), we observe almost no phase shift.

The adaptive algorithm adjusts the spatial mesh according to the "propagating wave" in the magnetic domain, as presented in Figure 5.9.3. Importantly, the spatial adaptive algorithm preserves the symmetry with respect to the midpoint of the domain. In Figure 5.9.4 we show the evolution of the time-step size, the selected BDF order and the spatial degrees of freedom. Until $t \approx 0.05$, the BDF(2) method is stable. After this time, the VSVO algorithm frequently tries to utilize higher-order BDF methods up to order 4 (and accordingly increases the time-steps due to the stability condition (4.5.5)), which are deemed unstable after several time-steps. This instability may be related to the small damping parameter $\alpha = 0.02$. For further details we refer to the discussion on small damping coefficients in Section 5.2 and Example 39. In total, we perform 16 863 time-steps with the linear solver requiring an average of 2.77 s computation time per step. As shown in Table 5.5.1, the linear solver accounts for the majority of the total computation time in the fully adaptive algorithm.

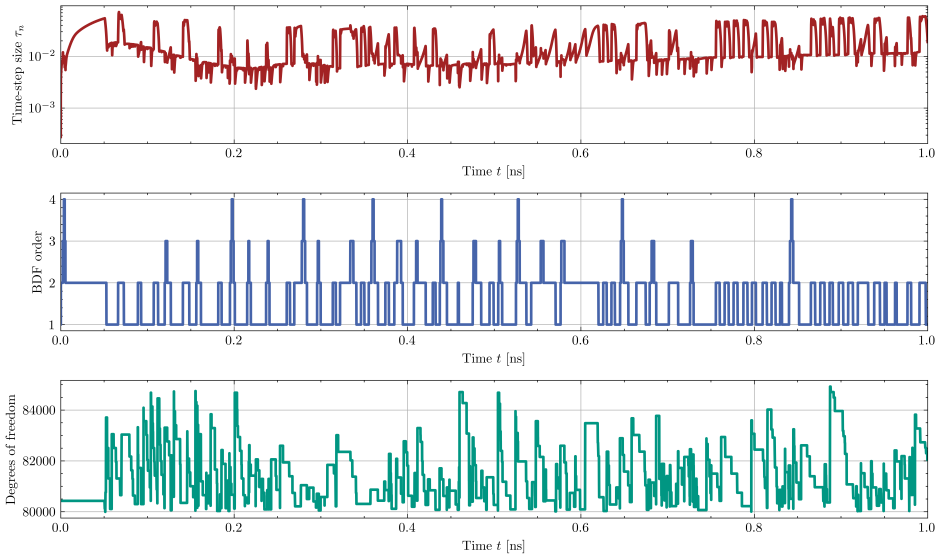


Figure 5.9.4: Evolution of the (dimensionless) time-step size (top), the selected BDF order (middle) and the spatial degrees of freedom (bottom) for the space and time adaptive algorithm using the BDF(1 - 4) method with linear finite elements.

Appendix

A.1 Polynomials

In this appendix, we provide the explicit definitions for the polynomials used to define the roots of the variable step-size BDF(4) polynomial (3.6.4). Specifically, we define

$$\begin{aligned}
\widetilde{a}(x,y,z) \coloneqq\ & x^4y^4z^2 + 6x^3y^4z^2 + 4x^3y^3z^2 + 2x^3y^3z + 15x^2y^4z^2 + 20x^2y^3z^2 + 10x^2y^3z \\
& + 6x^2y^2z^2 + 6x^2y^2z + x^2y^2 + 16xy^4z^2 + 32xy^3z^2 + 16xy^3z + 20xy^2z^2 + 20xy^2z \\
& + 4xy^2 + 4xyz^2 + 6xyz + 2xy + 6y^4z^2 + 16y^3z^2 + 8y^3z + 15y^2z^2 + 15y^2z + 3y^2 \\
& + 6yz^2 + 9yz + 3y + z^2 + 2z + 1, \\
\widetilde{c}(x,y,z) \coloneqq\ & x^2y^3z^3 + 2x^2y^2z^3 + x^2y^2z^2 + 2xy^3z^3 + 8xy^2z^3 + 4xy^2z^2 + 6xyz^3 + 6xyz^2 \\
& + 2xyz + y^3z^3 + 6y^2z^3 + 3y^2z^2 + 9yz^3 + 9yz^2 + 3yz + 4z^3 + 6z^2 + 4z + 1, \\
\widetilde{e}(x,y,z) \coloneqq\ & 4x^3y^2z + 9x^2y^2z + 6x^2yz + 3x^2y + 6xy^2z + 8xyz + 4xy \\
& + 2xz + 2x + y^2z + 2yz + y + z + 1,
\end{aligned}$$

for $x, y, z \in \mathbb{R}$.

A.2 Monotonicity of a rational function

In this section, we demonstrate that the rational function α defined in (3.7.9) is monotonically increasing for $x, y > 0$, cf. [92, Appendix A]. For convenience, we recall that

$$\alpha(x,y) = -\frac{b_1(x,y)}{b_0(x,y)}$$

for $x, y > 0$, where b_1 and b_0 are defined in (3.7.1). We begin by recalling the rational function b_1, which is given by

$$b_1(x,y) = -\frac{x^2}{1+x} - \frac{x^2y}{1+y+xy} - \frac{x^2y^2}{1+y+xy}\frac{1+x}{1+y}$$

for $x, y \geq 0$. By direct calculations, we obtain the first-order derivatives

$$\partial_x b_1(x,y) = \frac{-x\delta(x,y)}{(x+1)^2\,(y+1)\,(xy+y+1)^2}$$

and

$$\partial_y b_1(x,y) = -\frac{x^2\left((xy^2+2y^2+2y)(x+1)+(y+1)^2\right)}{(y+1)^2\,(xy+y+1)^2},$$

where

$$\delta(x,y) \coloneqq 2x^4y^3 + 10x^3y^3 + 5x^3y^2 + 20x^2y^3 + 20x^2y^2 + 4x^2y$$

$$+ 18xy^3 + 27xy^2 + 11xy + x + 6y^3 + 12y^2 + 8y + 2.$$

Similarly, we obtain for the rational function

$$b_0(x, y) = \frac{1 + 2x}{1 + x} + \frac{xy}{1 + y + xy}$$

the first-order derivatives

$$\begin{aligned}\partial_x b_0(x, y) &= \frac{1}{(1 + x)^2} + \frac{y + y^2}{(1 + y + xy)^2},\\ \partial_y b_0(x, y) &= \frac{x}{(1 + y + xy)^2}.\end{aligned}$$

Combining

$$\begin{aligned}b_1(x, y)\partial_x b_0(x, y) - \partial_x b_1(x, y)b_0(x, y) &= \frac{2x\left(3x^2y^2 + 6xy^2 + 3xy + 3y^2 + 3y + 1\right)}{x^2y^2 + x^2y + 2xy^2 + 3xy + x + y^2 + 2y + 1} > 0,\\ b_1(x, y)\partial_y b_0(x, y) - \partial_y b_1(x, y)b_0(x, y) &= \frac{x^2\left(3xy + 3y + 1\right)}{xy^3 + 2xy^2 + xy + y^3 + 3y^2 + 3y + 1} > 0\end{aligned}$$

for $x, y > 0$, we obtain for the first-order derivatives of α

$$\begin{aligned}\partial_x \alpha(x, y) &= \frac{b_1(x, y)\partial_x b_0(x, y) - \partial_x b_1(x, y)b_0(x, y)}{b_0(x, y)^2} > 0,\\ \partial_y \alpha(x, y) &= \frac{b_1(x, y)\partial_y b_0(x, y) - \partial_y b_1(x, y)b_0(x, y)}{b_0(x, y)^2} > 0\end{aligned}$$

since $b_0(x, y)^2 > 0$. Consequently, α is monotonically increasing for $x, y > 0$.

Abbreviations and symbols

In general, we denote vector-valued quantities by bold symbols.

Abbreviations

LLG	Landau–Lifshitz–Gilbert
VSVO	variable step-size variable order
BDF	backward differentiation formulas
BDF(k)	backward differentiation formulas of order k
BDF($k_{\min}$ - $k_{\max}$)	VSVO time-stepping with BDF($k_{\min}$) to BDF($k_{\max}$)
CFL	Courant–Friedrichs–Lewy
DC	discrete convolution
DOC	discrete orthogonal convolution
GMRES	generalized minimal residual
UMFPACK	Unsymmetric MultiFrontal PACKage
LU	lower-upper
ODE	ordinary differential equation

Physical quantities

$\boldsymbol{M}$	dimensional magnetization	A/m
M_s	saturation magnetization	A/m
$\boldsymbol{J}$	current density	A/m^2
A	exchange stiffness constant	J/m
μ_0	permeability of vacuum	N/A^2
γ_0	gyromagnetic ratio	m/(As)
α	Gilbert damping parameter	
$\mathcal{E}_{\text{ex}}$	exchange energy	N/m
$\mathcal{E}_{\text{ext}}$	external field energy	N/m
$\mathcal{E}_{\text{stray}}$	stray field energy	N/m
$\mathcal{E}_{\text{tot}}$	total magnetic Gibbs free energy	N/m
$\boldsymbol{H}_{\text{ext}}$	external field	A/m
$\boldsymbol{H}$	magnetic field	A/m
$\boldsymbol{H}_{\text{s}}$	stray field	A/m
$\boldsymbol{H}_{\text{eff}}$	effective field	A/m
U	scalar potential of the stray field	A

MATHEMATICAL NOTATION

$\mathbb{R}$	real numbers
$\mathbb{C}$	complex numbers
Ω	bounded domain in $\mathbb{R}^d$, $d \in \{2, 3\}$
$\partial\Omega$	boundary of Ω
T	final time
$\boldsymbol{x}$	spatial variable
t	time variable
Ω_T	time-space domain
L_{ref}	reference length
T_{ref}	reference time
$\boldsymbol{h}_{\mathrm{ext}}$	dimensionless external field
$\boldsymbol{h}_{\mathrm{s}}$	dimensionless stray field
$\boldsymbol{h}_{\mathrm{eff}}$	dimensionless effective field
$\boldsymbol{m}$	dimensionless magnetization
$\boldsymbol{v}$	time derivative of the dimensionless magnetization
$\boldsymbol{m}_0$	initial magnetization
ℓ_{ex}	exchange length
$\mathbb{S}^2$	unit sphere in $\mathbb{R}^3$
$\boldsymbol{T}(\boldsymbol{m})$	tangent space at $\boldsymbol{m}$
$\boldsymbol{P}(\boldsymbol{m})$	orthogonal projection onto $\boldsymbol{T}(\boldsymbol{m})$
$\boldsymbol{f}$	right-hand side
$\mathcal{C}^k(\Omega)$	space of k-times continuously differentiable functions
$\mathcal{C}_c^\infty(\Omega)$	space of infinitely differentiable functions with compact support
$\mathcal{L}^p(\Omega)$	Lebesgue space with $1 \leq p \leq \infty$
$\mathcal{L}^p(0, T; X)$	Bochner space with $1 \leq p \leq \infty$
$\mathcal{H}^k(0, T; X)$	k-times weakly differentiable functions in $\mathcal{L}^2(0, T; X)$
$\mathcal{W}^{k,p}(\Omega)$	Sobolev space with $k \in \mathbb{N}$ and $1 \leq p \leq \infty$
$\mathcal{H}^k(\Omega) = \mathcal{W}^{k,2}(\Omega)$	space of k-times weakly differentiable functions in $\mathcal{L}^2(\Omega)$ with $k \in \mathbb{N}$
$\mathcal{H}_0^1(\Omega)$	closure of $\mathcal{C}_c^\infty(\Omega)$ in $\mathcal{H}^1(\Omega)$
$\mathcal{H}^{-1}(\Omega)$	dual space of $\mathcal{H}_0^1(\Omega)$
$\widetilde{\mathcal{H}}^{-1}(\Omega)$	dual space of $\mathcal{H}^1(\Omega)$
$\mathcal{H}_\star^1(\Omega)$	$\mathcal{H}^1$-functions with zero Neumann boundary values
$\chi_\square$	indicator function of $\square$
$\square^{\mathrm{int}}$	restriction of $\square$ to Ω
$\square^{\mathrm{ext}}$	restriction of $\square$ to $\mathbb{R}^3 \setminus \overline{\Omega}$
$\lvert \cdot \rvert$	absolute value or Euclidean norm
$\lVert \cdot \rVert_\infty$	maximum norm or induced matrix norm
$\lVert \cdot \rVert_{\mathcal{L}^p}$	Lebesgue norm

$\|\cdot\| = \|\cdot\|_{\mathcal{L}^2}$	$\mathcal{L}^2$-norm
$\|\cdot\|_{\mathcal{W}^{k,p}}$	Sobolev norm
$\|\cdot\|_{\mathcal{L}^p(0,T;X)}$	Bochner norm
$(\cdot,\cdot)_\Omega = (\cdot,\cdot)$	inner product in $\mathcal{L}^2(\Omega)$
$\langle\cdot,\cdot\rangle$	duality pairing
$\|\cdot\|_{\mathcal{H}^{1/2}(\partial\Omega)}$	Slobodeckij semi-norm
$\mathcal{H}^{1/2}(\partial\Omega)$	fractional Sobolev space
$\mathcal{H}^{-1/2}(\partial\Omega)$	dual space of $\mathcal{H}^{1/2}(\partial\Omega)$
γ_{D}	Dirichlet trace operator
γ_{N}	Neumann trace operator
$\partial_{\boldsymbol{n}} u = \boldsymbol{\nabla} u \cdot \boldsymbol{n}$	normal derivative of u
$\boldsymbol{P}_0^n$	$\mathcal{L}^2$-projection onto $\boldsymbol{V}_h^n$
$\boldsymbol{I}$	identity operator
$a(\cdot,\cdot)$	bilinear form
$b^{\boldsymbol{m}}(\cdot,\cdot)$	bilinear form
τ_n	time-step size at t_n
$\mathcal{T}_h^n$	finite element triangulation at t_n
K	cell in $\mathcal{T}_h^n$
E	face or edge in Σ_n
h	maximal mesh size of $\mathcal{T}_h^n$
h_K	diameter of a cell $K \in \mathcal{T}_h^n$
h_E	length of edge E (in 2D) or diameter of face E (in 3D)
p	polynomial degree of finite element functions
k	BDF order
Σ_n	set of edges (2D) or faces (3D) of $\mathcal{T}_h^n$
Σ_n°	set of interior edges or faces in Σ_n
$\mathbb{P}_p$	space of polynomials of degree at most p
$\mathcal{S}_h^p(\mathcal{T}_h^n)$	space of globally continuous piecewise polynomials of degree p on $\mathcal{T}_h^n$
V	function space
V_h^n	finite element space
I_n	time interval
$\boldsymbol{T}_h^n(\boldsymbol{m})$	discrete tangent space at $\boldsymbol{m}$
$\boldsymbol{n}$	outer normal vector on $\partial\Omega$
$\boldsymbol{n}_E$	normal vector on E
$[\![\cdot]\!]_E$	jump across E
$\boldsymbol{m}_h^n$	discrete magnetization at t_n
$\boldsymbol{m}_h$	linear reconstruction of $\{\boldsymbol{m}_h^n\}_{n=0}^N$ in time
$\boldsymbol{M}_h$	three-point (quadratic) reconstruction of $\{\boldsymbol{m}_h^n\}_{n=0}^N$ in time
$\boldsymbol{w}$	linear time-space reconstruction of $\{\boldsymbol{m}_h^n\}_{n=0}^N$
$\boldsymbol{W}$	three-point (quadratic) time-space reconstruction of $\{\boldsymbol{m}_h^n\}_{n=0}^N$

λ_h^n	discrete Lagrangian multiplier at t_n
λ_h	linear reconstruction of $\{\lambda_h^n\}_{n=0}^N$ in time
$\widehat{\boldsymbol{m}}_h^n$	predictor for $\boldsymbol{m}_h^n$ at t_n
$\boldsymbol{v}_h^n$	discrete first time derivative of the magnetization at t_n
$\boldsymbol{f}_h^n$	right-hand side at t_n
$\boldsymbol{\varphi}_h^n$, ψ_h^n	finite element test functions at t_n
$\boldsymbol{\epsilon}$, $\boldsymbol{\mathcal{E}}$	error variable w.r.t linear and quadratic reconstructions of $\{\boldsymbol{m}_h^n\}_{n=0}^N$
$\widetilde{\boldsymbol{\epsilon}}$, $\widetilde{\boldsymbol{\mathcal{E}}}$	error variable w.r.t linear and quadratic time-space reconstructions of $\{\boldsymbol{m}_h^n\}_{n=0}^N$
ε	error variable for linear reconstruction of $\{\lambda_h^n\}_{n=0}^N$
$\boldsymbol{r}_h$	residual of the discrete solution
$\overline{\partial}^k(\cdot)$	discrete differential of order k
$\overline{\partial}_n^k(\cdot)$	projection of $\overline{\partial}^k(\cdot)$ onto V_h^n
$\overline{\partial}^B(\cdot)$	BDF(2) time discretization operator
$\overline{\partial}_n^B(\cdot)$	projection of $\overline{\partial}^B(\cdot)$ onto V_h^n
$-\Delta_h^n(\cdot)$	discrete Laplacian
$\boldsymbol{\mathcal{R}}\boldsymbol{m}_h^n$	elliptic reconstruction of $\boldsymbol{m}_h^n$
$R_K(\boldsymbol{m}_h^n)$	element (cell) residual on K
$R_E(\boldsymbol{m}_h^n)$	edge/face residual on E
$\eta(\boldsymbol{m}_h^n; X)$	elliptic a posteriori error estimator w.r.t space X
$\mathcal{E}_i$	time and reconstruction a posteriori error estimators for $i = 1, \ldots, 6$
$\mathcal{P}$	projection a posteriori error estimators
Λ_i	space a posteriori error estimators for $i = 1, 2, 3$
$\mathcal{F}_i$	data approximation a posteriori error estimators for $i = 1, 2$
Ξ_i	changing mesh a posteriori error estimators for $i = 1, 2, 3$
$\mathcal{C}_i$	finite element space conforming a posteriori error estimators for $i = 1, 2$
$\mathcal{Q}_i$	extrapolation a posteriori error estimators for $i = 1, 2$
$\mathcal{I}_i$	initial a posteriori error estimators for $i = 1, 2, 3$
$\boldsymbol{P}_h(\boldsymbol{m})$	orthogonal projection onto the discrete tangent space
u	arbitrary function $u : [0, T] \to \mathbb{R}$
u^n	u evaluated at t_n
$\widehat{u}^n$	predictor for u^n
f	arbitrary right-hand side
$\nabla_\tau^k(\cdot)$	discrete difference operator
δ_i	BDF coefficients
κ_n	consecutive time-step size ratio
$\kappa_{0,k}$	upper time-step ratio bounds for BDF(k)
$\alpha_{0,k}$	lower bounds for α for BDF(k)

Φ^k	general form of the BDF(k) method
$\tau_{j,\dots,\ell}$	sum of time-step sizes $\tau_j + \cdots + \tau_\ell$
L_j^n	Lagrange basis polynomials
$P[u\|t_0, \dots, t_{k-1}]$	interpolation polynomial
G	positive definite G-stability matrix
$\|\cdot\|_G$	norm corresponding to G
η_k	Nevanlinna–Odeh multiplier for BDF(k)
z_i	roots of the characteristic BDF polynomial
$b_j^{(k,n)}$	discrete convolution kernels
$\theta_{n-j}^{(k,n)}$	discrete orthogonal convolution kernels
$\widehat{\theta}_{m-\ell}^{(3,m)}$	scaled discrete orthogonal convolution kernels
$\widehat{\boldsymbol{\theta}}_{m-\ell}^{(3,m)}$	vector-valued scaled discrete orthogonal convolution
α_m, β_m	rational functions
T_k	leading term of the local truncation error of BDF(k)
$\overline{\partial}_\tau^k(\cdot)$	finite-difference approximation of the k-th time derivative
$\underline{\kappa}_k$	implemented lower bound for time-step ratio for BDF(k)
$\overline{\kappa}_k$	implemented upper bound for time-step ratio for BDF(k)
$\tau_{\min}$	lower bound on the time-step size
$\tau_{\max}$	upper bound on the time-step size
$\tau_n^{(k)}$	estimated time-step size for BDF(k)
$\boldsymbol{P}^\ell[\boldsymbol{m}_h^{n-\ell}, \dots, \boldsymbol{m}_h^n]$	interpolation polynomial of order ℓ for $\{\boldsymbol{m}_h^j\}_{j=n-\ell}^n$
$\boldsymbol{m}_h[t_n, \dots, t_{n-j}]$	divided differences
$\boldsymbol{G}_h$	gradient recovery operator
θ_{r}	mesh refinement parameter
θ_{c}	mesh coarsening parameter
$k_{\min}$, $k_{\max}$	lower and upper bound on the BDF order
$\mathrm{TOL}_{\mathrm{t}}$	prescribed time tolerance
$\mathrm{TOL}_{\mathrm{s}}$	prescribed space tolerance
$\eta_K^\star$	local spatial error indicator for $\star \in \{\Sigma, \mathcal{T}, \boldsymbol{G}\}$
$\eta^\star$	global spatial error indicator for $\star \in \{\Sigma, \mathcal{T}, \boldsymbol{G}\}$

Bibliography

[1] Abert, C., Hrkac, G., Page, M., Praetorius, D., Ruggeri, M., and Suess, D. Spin-polarized transport in ferromagnetic multilayers: an unconditionally convergent FEM integrator. *Comput. Math. Appl.*, 68(6):639–654, 2014. doi:10.1016/j.camwa.2014.07.010.

[2] Adams, R. A. and Fournier, J. J. F. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003. ISBN 0-12-044143-8.

[3] Aharoni, A. *Introduction to the Theory of Ferromagnetism*, volume 109. Clarendon Press, 2000.

[4] Ainsworth, M. and Oden, J. T. A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Engrg.*, 142(1-2):1–88, 1997. doi:10.1016/S0045-7825(96)01107-3.

[5] Ainsworth, M. and Oden, J. T. *A posteriori error estimation in finite element analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000. ISBN 0-471-29411-X. doi:10.1002/9781118032824.

[6] Akrivis, G. and Chatzipantelidis, P. A posteriori error estimates for the two-step backward differentiation formula method for parabolic equations. *SIAM J. Numer. Anal.*, 48(1):109–132, 2010. doi:10.1137/090756995.

[7] Akrivis, G., Feischl, M., Kovács, B., and Lubich, C. Higher-order linearly implicit full discretization of the Landau-Lifshitz-Gilbert equation. *Math. Comp.*, 90(329):995–1038, 2021. doi:10.1090/mcom/3597.

[8] Akrivis, G., Chen, M., Yu, F., and Zhou, Z. The energy technique for the six-step BDF method. *SIAM J. Numer. Anal.*, 59(5):2449–2472, 2021. doi:10.1137/21M1392656.

[9] Akrivis, G., Chen, M., Han, J., Yu, F., and Zhang, Z. The variable two-step BDF method for parabolic equations. *BIT*, 64(1):Paper No. 14, 21, 2024. doi:10.1007/s10543-024-01007-y.

[10] Akrivis, G., Bartels, S., and Palus, C. Quadratic constraint consistency in the projection-free approximation of harmonic maps and bending isometries. *Math. Comp.*, 94(355): 2251–2269, 2025. doi:10.1090/mcom/4035.

[11] Akrivis, G., Bartels, S., Ruggeri, M., and Wang, J. Projection-free approximation of flows of harmonic maps with quadratic constraint accuracy and variable step sizes, 2025. URL `https://arxiv.org/abs/2505.05655`.

[12] Aldé, M., Feischl, M., and Praetorius, D. BDF2-type integrator for Landau-Lifshitz-Gilbert equation in micromagnetics: unconditional weak convergence to weak solutions, 2026. URL `https://arxiv.org/abs/2511.22000`.

[13] Alouges, F. and Jaisson, P. Convergence of a finite element discretization for the Landau-Lifshitz equations in micromagnetism. *Math. Models Methods Appl. Sci.*, 16(2): 299–316, 2006. doi:10.1142/S0218202506001169.

[14] Alouges, F. and Soyeur, A. On global weak solutions for Landau-Lifshitz equations: existence and nonuniqueness. *Nonlinear Anal.*, 18(11):1071–1084, 1992. doi:10.1016/0362-546X(92)90196-L.

[15] Alouges, F., Kritsikis, E., Steiner, J., and Toussaint, J.-C. A convergent and precise finite element scheme for Landau-Lifschitz-Gilbert equation. *Numer. Math.*, 128(3): 407–430, 2014. doi:10.1007/s00211-014-0615-3.

[16] Alouges, F. A new finite element scheme for Landau-Lifchitz equations. *Discrete Contin. Dyn. Syst. Ser. S*, 1(2):187–196, 2008. doi:10.3934/dcdss.2008.1.187.

[17] An, R., Li, Y., and Sun, W. Optimal error analysis of the normalized tangent plane FEM for Landau-Lifshitz-Gilbert equation. *IMA J. Numer. Anal.*, 45(5):3109–3137, 2025. doi:10.1093/imanum/drae084.

[18] Arndt, D., Bangerth, W., Bergbauer, M., Feder, M., Fehling, M., Heinz, J., Heister, T., Heltai, L., Kronbichler, M., Maier, M., Munch, P., Pelteret, J.-P., Turcksin, B., Wells, D., and Zampini, S. The deal.II Library, Version 9.5. *Journal of Numerical Mathematics*, 31(3):231–246, 2023. doi:10.1515/jnma-2023-0089.

[19] Bagneres-Viallix, A., Baras, P., and Albertini, J. 2D and 3D calculations of micromagnetic wall structures using finite elements. *IEEE Transactions on Magnetics*, 27(5): 3819–3822, 1991. doi:10.1109/20.104934.

[20] Baibich, M. N., Broto, J. M., Fert, A., Van Dau, F. N., Petroff, F., Etienne, P., Creuzet, G., Friederich, A., and Chazelas, J. Giant Magnetoresistance of (001)Fe/(001)Cr Magnetic Superlattices. *Phys. Rev. Lett.*, 61:2472–2475, Nov 1988. doi:10.1103/PhysRevLett.61.2472.

[21] Baňas, L. and Slodička, M. Error estimates for Landau-Lifshitz-Gilbert equation with magnetostriction. *Appl. Numer. Math.*, 56(8):1019–1039, 2006. doi:10.1016/j.apnum.2005.09.003.

[22] Baňas, L., Bartels, S., and Prohl, A. A convergent implicit finite element discretization of the Maxwell-Landau-Lifshitz-Gilbert equation. *SIAM J. Numer. Anal.*, 46(3):1399–1422, 2008. doi:10.1137/070683064.

[23] Baňas, L. *On dynamical Micromagnetism with Magnetostriction.* PhD thesis, Ghent University, 2005.

[24] Baňas, L. Adaptive techniques for Landau-Lifshitz-Gilbert equation with magnetostriction. *J. Comput. Appl. Math.*, 215(2):304–310, 2008. doi:10.1016/j.cam.2006.03.043.

[25] Bänsch, E. and Brenner, A. A posteriori estimates for the two-step backward differentiation formula and discrete regularity for the time-dependent Stokes equations. *IMA J. Numer. Anal.*, 39(2):713–759, 2019. doi:10.1093/imanum/dry014.

[26] Bänsch, E., Karakatsani, F., and Makridakis, C. G. The effect of mesh modification in time on the error control of fully discrete approximations for parabolic equations. *Appl. Numer. Math.*, 67:35–63, 2013. doi:10.1016/j.apnum.2011.08.008.

[27] Bänsch, E. Local mesh refinement in 2 and 3 dimensions. *Impact Comput. Sci. Engrg.*, 3(3):181–191, 1991. doi:10.1016/0899-8248(91)90006-G.

[28] Bartels, S. and Prohl, A. Convergence of an implicit finite element method for the Landau-Lifshitz-Gilbert equation. *SIAM J. Numer. Anal.*, 44(4):1405–1419, 2006. doi:10.1137/050631070.

[29] Bartels, S., Ko, J., and Prohl, A. Numerical analysis of an explicit approximation scheme for the Landau-Lifshitz-Gilbert equation. *Math. Comp.*, 77(262):773–788, 2008. doi:10.1090/S0025-5718-07-02079-0.

[30] Bartels, S., Kovács, B., and Wang, Z. Error analysis for the numerical approximation of the harmonic map heat flow with nodal constraints. *IMA J. Numer. Anal.*, 44(2): 633–653, 2024. doi:10.1093/imanum/drad037.

[31] Bartels, S. Projection-free approximation of geometrically constrained partial differential equations. *Math. Comp.*, 85(299):1033–1049, 2016. doi:10.1090/mcom/3008.

[32] Becker, J. A second order backward difference method with variable steps for a parabolic problem. *BIT*, 38(4):644–662, 1998. doi:10.1007/BF02510406.

[33] Bethuel, F., Coron, J.-M., Ghidaglia, J.-M., and Soyeur, A. Heat flows and relaxed energies for harmonic maps. In *Nonlinear diffusion equations and their equilibrium states, 3 (Gregynog, 1989)*, volume 7 of *Progr. Nonlinear Differential Equations Appl.*, pages 99–109. Birkhäuser Boston, Boston, MA, 1992. ISBN 0-8176-3531-9. doi:10.1007/978-1-4612-0393-3_7.

[34] Binasch, G., Grünberg, P., Saurenbach, F., and Zinn, W. Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange. *Phys. Rev. B*, 39:4828–4830, Mar 1989. doi:10.1103/PhysRevB.39.4828.

[35] Bjørk, R., Poulsen, E., Nielsen, K., and Insinga, A. MagTense: A micromagnetic framework using the analytical demagnetization tensor. *Journal of Magnetism and Magnetic Materials*, 535:168057, 2021. doi:10.1016/j.jmmm.2021.168057.

[36] Bohn, J., Dörfler, W., Feischl, M., and Karch, S. Adaptive mesh refinement for the Landau–Lifshitz–Gilbert equation, 2026. URL `https://arxiv.org/abs/2303.07463`.

[37] Brent, R. P. *Algorithms for minimization without derivatives.* Prentice-Hall Series in Automatic Computation. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1973.

[38] Bruckner, F., Suess, D., Feischl, M., Führer, T., Goldenits, P., Page, M., Praetorius, D., and Ruggeri, M. Multiscale modeling in micromagnetics: existence of solutions and numerical integration. *Math. Models Methods Appl. Sci.*, 24(13):2627–2662, 2014. doi:10.1142/S0218202514500328.

[39] Calvo, M., Grande, T., and Grigorieff, R. D. On the zero stability of the variable order variable stepsize BDF-formulas. *Numer. Math.*, 57(1):39–50, 1990. doi:10.1007/BF01386395.

[40] Carbou, G. and Fabrie, P. Regular solutions for Landau-Lifschitz equation in a bounded domain. *Differential Integral Equations*, 14(2):213–229, 2001.

[41] Carbou, G. and Fabrie, P. Regular solutions for Landau-Lifschitz equation in $\mathbb{R}^3$. *Commun. Appl. Anal.*, 5(1):17–30, 2001.

[42] Carstensen, C. and Verfürth, R. Edge residuals dominate a posteriori error estimates for low order finite element methods. *SIAM J. Numer. Anal.*, 36(5):1571–1587, 1999. doi:10.1137/S003614299732334X.

[43] Celaya, E. A., Aguirrezabala, J. J. A., and Chatzipantelidis, P. Implementation of an Adaptive BDF2 Formula and Comparison with the MATLAB Ode15s. *Procedia Computer Science*, 29:1014–1026, 2014. doi:10.1016/j.procs.2014.05.091. 2014 International Conference on Computational Science.

[44] Cheng, Q. and Shen, J. Length preserving numerical schemes for Landau-Lifshitz equation based on Lagrange multiplier approaches. *SIAM J. Sci. Comput.*, 45(2): A530–A553, 2023. doi:10.1137/22M1501143.

[45] Cimrák, I. A survey on the numerics and computations for the Landau-Lifshitz equation of micromagnetism. *Arch. Comput. Methods Eng.*, 15(3):277–309, 2008. doi:10.1007/s11831-008-9021-2.

[46] Contri, A., Kovács, B., and Massing, A. Error Analysis of BDF 1–6 Time-Stepping Methods for the Transient Stokes Problem: Velocity and Pressure Estimates. *SIAM J. Numer. Anal.*, 63(4):1586–1616, 2025. doi:10.1137/23M1606800.

[47] Coron, J.-M. Nonuniqueness for the heat flow of harmonic maps. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 7(4):335–344, 1990. doi:10.1016/S0294-1449(16)30295-5.

[48] Crouzeix, M. and Lisbona, F. J. The convergence of variable-stepsize, variable-formula, multistep methods. *SIAM J. Numer. Anal.*, 21(3):512–534, 1984. doi:10.1137/0721037.

[49] Dahlquist, G. G-stability is equivalent to A-stability. *BIT*, 18(4):384–401, 1978. doi:10.1007/BF01932018.

[50] Davis, T. A. Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software*, 30(2):196–199, 2004. doi:10.1145/992200.992206.

[51] Frutos, J. de and Novo, J. A posteriori error estimation with the p-version of the finite element method for nonlinear parabolic differential equations. *Comput. Methods Appl. Mech. Engrg.*, 191(43):4893–4904, 2002. doi:10.1016/S0045-7825(02)00419-X.

[52] Frutos, J. de and Novo, J. Element-wise a posteriori estimates based on hierarchical bases for non-linear parabolic problems. *Internat. J. Numer. Methods Engrg.*, 63(8): 1146–1173, 2005. doi:10.1002/nme.1310.

[53] DeCaria, V., Guzel, A., Layton, W., and Li, Y. A variable stepsize, variable order family of low complexity. *SIAM J. Sci. Comput.*, 43(3):A2130–A2160, 2021. doi:10.1137/19M1258153.

[54] Demlow, A. Localized pointwise a posteriori error estimates for gradients of piecewise linear finite element approximations to second-order quasilinear elliptic problems. *SIAM J. Numer. Anal.*, 44(2):494–514, 2006. doi:10.1137/040610064.

[55] Deuflhard, P. and Bornemann, F. *Numerische Mathematik 2: Gewöhnliche Differentialgleichungen*. de Gruyter, Berlin, 2 edition, 2002.

[56] Deuflhard, P. and Weiser, M. *Adaptive numerical solution of PDEs*. De Gruyter Textbook. Walter de Gruyter & Co., Berlin, 2012. ISBN 978-3-11-028310-5. doi:10.1515/9783110283112.

[57] Di Fratta, G., Innerberger, M., and Praetorius, D. Weak-strong uniqueness for the Landau-Lifshitz-Gilbert equation in micromagnetics. *Nonlinear Anal. Real World Appl.*, 55:103122, 13, 2020. doi:10.1016/j.nonrwa.2020.103122.

[58] Di Fratta, G., Pfeiler, C.-M., Praetorius, D., Ruggeri, M., and Stiftner, B. Linear second-order IMEX-type integrator for the (eddy current) Landau-Lifshitz-Gilbert equation. *IMA J. Numer. Anal.*, 40(4):2802–2838, 2020. doi:10.1093/imanum/drz046.

[59] Dörfler, W. A time- and space-adaptive algorithm for the linear time-dependent Schrödinger equation. *Numer. Math.*, 73(4):419–448, 1996. doi:10.1007/s002110050199.

[60] Emmrich, E. Stability and error of the variable two-step BDF for semilinear parabolic problems. *J. Appl. Math. Comput.*, 19(1-2):33–55, 2005. doi:10.1007/BF02935787.

[61] Evans, L. C. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998. ISBN 0-8218-0772-2. doi:10.1090/gsm/019.

[62] Exl, L., Mauser, N. J., Schrefl, T., and Suess, D. The extrapolated explicit midpoint scheme for variable order and step size controlled integration of the Landau-Lifschitz-Gilbert equation. *J. Comput. Phys.*, 346:14–24, 2017. doi:10.1016/j.jcp.2017.06.005.

[63] Fang, Z. and Wang, X. An adaptive moving mesh method for simulating finite-time blowup solutions of the Landau-Lifshitz-Gilbert equation. *East Asian J. Appl. Math.*, 14(3):601–635, 2024.

[64] Feischl, M. and Tran, T. 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. doi:10.1137/16M1103427.

[65] Feischl, M. and Tran, T. The eddy current–LLG equations: FEM-BEM coupling and a priori error estimates. *SIAM J. Numer. Anal.*, 55(4):1786–1819, 2017. doi:10.1137/16M1065161.

[66] Fert, A., Cros, V., and Sampaio, J. Skyrmions on the track. *Nat. Nanotechnol.*, 8: 152–156, 2013.

[67] Fert, A., Reyren, N., and Cros, V. Magnetic skyrmions: advances in physics and potential applications. *Nature Reviews Materials*, 2(7):17031, June 2017. doi:10.1038/natrevmats.2017.31.

[68] Fert, A., Ramesh, R., Garcia, V., Casanova, F., and Bibes, M. Electrical control of magnetism by electric field and current-induced torques. *Rev. Mod. Phys.*, 96:015005, Mar 2024. doi:10.1103/RevModPhys.96.015005.

[69] Fert, A. Nobel lecture: Origin, development, and future of spintronics. *Rev. Mod. Phys.*, 80:1517–1530, Dec 2008. doi:10.1103/RevModPhys.80.1517.

[70] Fredkin, D. and Koehler, T. Numerical micromagnetics by the finite element method. *IEEE Transactions on Magnetics*, 23(5):3385–3387, 1987. doi:10.1109/TMAG.1987.1065578.

[71] Fredkin, D. and Koehler, T. Hybrid method for computing demagnetizing fields. *IEEE Transactions on Magnetics*, 26(2):415–417, 1990. doi:10.1109/20.106342.

[72] García-Cervera, C. J. and Roma, A. M. Adaptive mesh refinement for micromagnetics simulations. *IEEE transactions on magnetics*, 42(6):1648–1654, 2006.

[73] Gear, C. W. and Watanabe, D. S. Stability and convergence of variable order multistep methods. *SIAM J. Numer. Anal.*, 11:1044–1058, 1974. doi:10.1137/0711080.

[74] Gilbert, T. L. and Kelly, J. M. Anomalous rotational damping in ferromagnetic sheets. In *Conference on Magnetism and Magnetic Materials*, pages 253–263, Pittsburgh, PA, June 1955. American Institute of Electrical Engineers. Published Oct. 1955.

[75] Gilbert, T. A phenomenological theory of damping in ferromagnetic materials. *IEEE Transactions on Magnetics*, 40(6):3443–3449, 2004. doi:10.1109/TMAG.2004.836740.

[76] Goldenits, P. *Konvergente numerische Integration der Landau-Lifshitz-Gilbert Gleichung.* PhD thesis, TU Wien, Institute for Analysis and Scientific Computing, Vienna University of Technology, 11 2012. URL `https://resolver.obvsg.at/urn:nbn:at:at-ubtuw:1-49520`.

[77] Grigorieff, R. D. Stability of multistep-methods on variable grids. *Numer. Math.*, 42(3): 359–377, 1983. doi:10.1007/BF01389580.

[78] Guglielmi, N. and Zennaro, M. On the zero-stability of variable stepsize multistep methods: the spectral radius approach. *Numer. Math.*, 88(3):445–458, 2001. doi:10.1007/s211-001-8010-0.

[79] Guo, B. L. and Hong, M. C. The Landau-Lifshitz equation of the ferromagnetic spin chain and harmonic maps. *Calc. Var. Partial Differential Equations*, 1(3):311–334, 1993. doi:10.1007/BF01191298.

[80] Hairer, E. and Wanner, G. *Solving ordinary differential equations. II*, volume 14 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1996. ISBN 3-540-60452-9. doi:10.1007/978-3-642-05221-7. Stiff and differential-algebraic problems.

[81] Hairer, E., Nørsett, S. P., and Wanner, G. *Solving ordinary differential equations. I*, volume 8 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 1993. ISBN 3-540-56670-8. Nonstiff problems.

[82] Hay, A., Etienne, S., Pelletier, D., and Garon, A. hp-adaptive time integration based on the BDF for viscous flows. *J. Comput. Phys.*, 291:151–176, 2015. doi:10.1016/j.jcp.2015.03.022.

[83] Hay, A., Etienne, S., Garon, A., and Pelletier, D. Time-integration for ALE simulations of fluid-structure interaction problems: stepsize and order selection based on the BDF. *Comput. Methods Appl. Mech. Engrg.*, 295:172–195, 2015. doi:10.1016/j.cma.2015.06.006.

[84] Hertel, R. and Kronmuller, H. Adaptive finite element mesh refinement techniques in three-dimensional micromagnetic modeling. *IEEE Transactions on Magnetics*, 34(6): 3922–3930, 1998. doi:10.1109/20.728305.

[85] Hirohata, A., Yamada, K., Nakatani, Y., Prejbeanu, I.-L., Diény, B., Pirro, P., and Hillebrands, B. Review on spintronics: Principles and device applications. *Journal of Magnetism and Magnetic Materials*, 509:166711, 2020. doi:10.1016/j.jmmm.2020.166711.

[86] Hrkac, G., Pfeiler, C.-M., Praetorius, D., Ruggeri, M., Segatti, A., and Stiftner, B. Convergent tangent plane integrators for the simulation of chiral magnetic skyrmion dynamics. *Adv. Comput. Math.*, 45(3):1329–1368, 2019. doi:10.1007/s10444-019-09667-z.

[87] Hubert, A. and Schäfer, R. *Magnetic domains: the analysis of magnetic microstructures.* Springer Science & Business Media, 1998.

[88] Kreuzer, C., Möller, C. A., Schmidt, A., and Siebert, K. G. Design and convergence analysis for an adaptive discretization of the heat equation. *IMA J. Numer. Anal.*, 32 (4):1375–1403, 2012. doi:10.1093/imanum/drr026.

[89] Kružík, M. and Prohl, A. Recent developments in the modeling, analysis, and numerics of ferromagnetism. *SIAM Rev.*, 48(3):439–483, 2006. doi:10.1137/S0036144504446187.

[90] Kunert, G. and Verfürth, R. Edge residuals dominate a posteriori error estimates for linear finite element methods on anisotropic triangular and tetrahedral meshes. *Numer. Math.*, 86(2):283–303, 2000. doi:10.1007/PL00005407.

[91] Lakkis, O. and Makridakis, C. G. Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems. *Math. Comp.*, 75(256):1627–1658, 2006. doi:10.1090/S0025-5718-06-01858-8.

[92] Li, Z. and Liao, H.-L. Stability of variable-step BDF2 and BDF3 methods. *SIAM J. Numer. Anal.*, 60(4):2253–2272, 2022. doi:10.1137/21M1462398.

[93] Liao, H.-l. and Zhang, Z. Analysis of adaptive BDF2 scheme for diffusion equations. *Math. Comp.*, 90(329):1207–1226, 2021. doi:10.1090/mcom/3585.

[94] Liao, H.-l., Tang, T., and Zhou, T. Discrete energy analysis of the third-order variable-step BDF time-stepping for diffusion equations. *J. Comput. Math.*, 41(2):325–344, 2023. doi:10.4208/jcm.2207-m2022-0020.

[95] Lifshitz, E. and Landau, L. On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. *Phys. Z. Sowjetunion*, 8(135), 1935. doi:10.1016/B978-0-08-036364-6.50008-9.

[96] Lozinski, A., Picasso, M., and Prachittham, V. An anisotropic error estimator for the Crank-Nicolson method: application to a parabolic problem. *SIAM J. Sci. Comput.*, 31 (4):2757–2783, 2009. doi:10.1137/080715135.

[97] Makridakis, C. and Nochetto, R. H. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, 41(4):1585–1594, 2003. doi:10.1137/S0036142902406314.

[98] McLean, W. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, Cambridge, 2000. ISBN 0-521-66332-6; 0-521-66375-X.

[99] NIST Micromagnetic Modeling Activity Group (μMAG) website. URL `https://www.ctcms.nist.gov/~rdm/mumag.html`. Accessed on February 10, 2026.

[100] Nevanlinna, O. and Odeh, F. Multiplier techniques for linear multistep methods. *Numer. Funct. Anal. Optim.*, 3(4):377–423, 1981. doi:10.1080/01630568108816097.

[101] Nochetto, R. H., Schmidt, A., Siebert, K. G., and Veeser, A. Pointwise a posteriori error estimates for monotone semi-linear equations. *Numer. Math.*, 104(4):515–538, 2006. doi:10.1007/s00211-006-0027-0.

[102] Normington, H. and Ruggeri, M. Convergent finite element methods for antiferromagnetic and ferrimagnetic materials. *ESAIM Math. Model. Numer. Anal.*, 59(1):167–199, 2025. doi:10.1051/m2an/2024065.

[103] Pfeiler, C.-M., Ruggeri, M., Stiftner, B., Exl, L., Hochsteger, M., Hrkac, G., Schöberl, J., Mauser, N. J., and Praetorius, D. Computational micromagnetics with Commics. *Computer Physics Communications*, 248:106965, March 2020. doi:10.1016/j.cpc.2019.106965.

[104] Plato, R. *Numerische Mathematik kompakt*. Friedr. Vieweg & Sohn, Wiesbaden, second edition, 2004. ISBN 3-528-13153-5. doi:10.1007/978-3-322-93922-7. Grundlagenwissen für Studium und Praxis. [Foundations for study and practice].

[105] Praetorius, D., Ruggeri, M., and Stiftner, B. Convergence of an implicit-explicit midpoint scheme for computational micromagnetics. *Comput. Math. Appl.*, 75(5):1719–1738, 2018. doi:10.1016/j.camwa.2017.11.028.

[106] Prohl, A. *Computational micromagnetism*. Advances in Numerical Mathematics. B. G. Teubner, Stuttgart, 2001. ISBN 3-519-00358-9. doi:10.1007/978-3-663-09498-2.

[107] Ruggeri, M. *Coupling and numerical integration of the Landau–Lifshitz–Gilbert equation*. PhD thesis, TU Wien, 2016. doi:10.34726/hss.2016.40261.

[108] Saad, Y. and Schultz, M. H. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, 1986. doi:10.1137/0907058.

[109] Saad, Y. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003. ISBN 0-89871-534-2. doi:10.1137/1.9780898718003.

[110] Sauter, S. A. and Schwab, C. *Boundary element methods*, volume 39 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2011. ISBN 978-3-540-68092-5. doi:10.1007/978-3-540-68093-2. Translated and expanded from the 2004 German original.

[111] Scholz, W., Schrefl, T., and Fidler, J. Mesh refinement in FE-micromagnetics for multi-domain Nd2Fe14B particles. *Journal of Magnetism and Magnetic Materials*, 196-197: 933–934, 1999. doi:10.1016/S0304-8853(98)00994-9.

[112] Schrefl, T. Finite elements in numerical micromagnetics. Part I: Granular hard magnets. *Journal of Magnetism and Magnetic Materials*, 207(1):45–65, 1999. doi:10.1016/S0304-8853(99)00532-6.

[113] Shampine, L. F. and Reichelt, M. W. The MATLAB ODE suite. volume 18, pages 1–22. 1997. doi:10.1137/S1064827594276424. Dedicated to C. William Gear on the occasion of his 60th birthday.

[114] Shepherd, D. *Numerical methods for dynamic micromagnetics*. PhD thesis, University of Manchester, 2015. URL `https://research.manchester.ac.uk/en/studentTheses/numerical-methods-for-dynamic-micromagnetics`.

[115] Söderlind, G., Fekete, I., and Faragó, I. On the zero-stability of multistep methods on smooth nonuniform grids. *BIT*, 58(4):1125–1143, 2018. doi:10.1007/s10543-018-0716-y.

[116] Tako, K., Schrefl, T., Wongsam, M., and Chantrell, R. Finite element micromagnetic simulations with adaptive mesh refinement. *Journal of applied physics*, 81(8):4082–4084, 1997.

[117] Vansteenkiste, A. and Van de Wiele, B. MuMax: A new high-performance micromagnetic simulation tool. *Journal of Magnetism and Magnetic Materials*, 323(21):2585–2591, 2011. doi:10.1016/j.jmmm.2011.05.037.

[118] Verfürth, R. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Advances in numerical mathematics. Wiley, 1996. ISBN 978-0-471-96795-8.

[119] Verfürth, R. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013. ISBN 978-0-19-967942-3. doi:10.1093/acprof:oso/9780199679423.001.0001.

[120] Visintin, A. On Landau-Lifshitz' equations for ferromagnetism. *Japan J. Appl. Math.*, 2 (1):69–84, 1985. doi:10.1007/BF03167039.

[121] Wang, D. and Ruuth, S. J. Variable step-size implicit-explicit linear multistep methods for time-dependent partial differential equations. *J. Comput. Math.*, 26(6):838–855, 2008.

[122] Yu, D. H. Asymptotically exact a posteriori error estimator for elements of bi-even degree. *Chinese J. Numer. Math. Appl.*, 13(2):64–78, 1991.

[123] Yu, D. H. Asymptotically exact a posteriori error estimators for elements of bi-odd degree. *Chinese J. Numer. Math. Appl.*, 13(4):82–90, 1991.

[124] Žutić, I., Fabian, J., and Das Sarma, S. Spintronics: Fundamentals and applications. *Rev. Mod. Phys.*, 76:323–410, Apr 2004. doi:10.1103/RevModPhys.76.323.