

ITERATIVE SOLUTION AND PRECONDITIONING FOR THE TANGENT PLANE SCHEME IN COMPUTATIONAL MICROMAGNETICS

JOHANNES KRAUS, CARL-MARTIN PFEILER, DIRK PRAETORIUS, MICHELE RUGGERI,
AND BERNHARD STIFTNER

ABSTRACT. The tangent plane scheme is a time-marching scheme for the numerical solution of the nonlinear parabolic Landau–Lifshitz–Gilbert equation (LLG), which describes the time evolution of ferromagnetic configurations. Exploiting the geometric structure of LLG, the tangent plane scheme requires only the solution of one linear variational form per time-step, which is posed in the discrete tangent space determined by the nodal values of the current magnetization. We develop an effective solution strategy for the arising constrained linear systems, which is based on appropriate Householder reflections. We derive possible preconditioners, which are (essentially) independent of the time-step, and prove that the preconditioned GMRES algorithm leads to linear convergence. Numerical experiments underpin the theoretical findings.

1. INTRODUCTION

1.1. Landau–Lifshitz–Gilbert equation. The Landau–Lifshitz–Gilbert equation (LLG) describes time-dependent micromagnetic phenomena in a ferromagnetic domain $\Omega \subset \mathbb{R}^3$ being bounded and Lipschitz with boundary $\partial\Omega$ [Gil55, LL08]. It reads

$$\partial_t \mathbf{m} = -\mathbf{m} \times \mathbf{h}_{\text{eff}}(\mathbf{m}) + \alpha \mathbf{m} \times \partial_t \mathbf{m} \quad \text{in } (0, T) \times \Omega, \quad (1a)$$

$$\partial_n \mathbf{m} = \mathbf{0} \quad \text{on } (0, T) \times \partial\Omega, \quad (1b)$$

$$\mathbf{m}(0) = \mathbf{m}^0 \quad \text{in } \Omega, \quad (1c)$$

where the unknown $\mathbf{m} : (0, T) \times \Omega \rightarrow \mathbb{R}^3$ is the magnetization, $\mathbf{h}_{\text{eff}}(\mathbf{m})$ is the \mathbf{m} -dependent effective field, $\alpha \in (0, 1]$ is the Gilbert damping constant, $T > 0$ is the final time, and $\mathbf{m}^0 \in \mathbf{H}^1(\Omega) := (H^1(\Omega))^3$ with $|\mathbf{m}^0| = 1$ a.e. in Ω is the initial configuration. The effective field comprises several contributions, which correspond to different phenomena in micromagnetism. In usual applications (cf., e.g., [HS98]), one has $\mathbf{h}_{\text{eff}}(\mathbf{m}) := \ell_{\text{ex}}^2 \Delta \mathbf{m} + \boldsymbol{\pi}(\mathbf{m}) + \mathbf{f}$, where $\ell_{\text{ex}}^2 \Delta \mathbf{m}$ is the exchange contribution with the exchange length $\ell_{\text{ex}} > 0$, $\boldsymbol{\pi} : \mathbf{H}^1(\Omega) \rightarrow \mathbf{L}^2(\Omega) := (L^2(\Omega))^3$ is a bounded operator, which collects all \mathbf{m} -dependent lower-order terms such as the stray field or the magnetocrystalline anisotropy contribution, $\mathbf{f} \in C^1([0, T], \mathbf{L}^2(\Omega))$ is the applied external field.

Taking the scalar product with \mathbf{m} in (1a), we note the PDE inherent constraints

$$\frac{1}{2} \partial_t |\mathbf{m}|^2 = \mathbf{m} \cdot \partial_t \mathbf{m} = 0 \quad \text{and thus} \quad |\mathbf{m}| = 1 \quad \text{a.e. in } (0, T) \times \Omega. \quad (2)$$

In particular, $\partial_t \mathbf{m}(t)$ belongs to the tangent space of $\mathbf{m}(t)$ for almost all $t \in (0, T)$. Using vector identities and (2), one can prove that (1a) is equivalent to

$$\alpha \partial_t \mathbf{m} + \mathbf{m} \times \partial_t \mathbf{m} = \mathbf{h}_{\text{eff}}(\mathbf{m}) - (\mathbf{h}_{\text{eff}}(\mathbf{m}) \cdot \mathbf{m}) \mathbf{m}, \quad (3)$$

2010 *Mathematics Subject Classification.* 35K35, 65M60, 65F08, 65F10, 65M22, 65Z05.

Key words and phrases. micromagnetics, finite elements, tangent plane scheme, preconditioning.

Acknowledgements. The authors acknowledge support from the Vienna Science and Technology Fund WWTF (grant MA14-44) and the Austrian Science Fund FWF (grants DK W1245 and SFB F65).

which is a linear equation in $\mathbf{v} := \partial_t \mathbf{m}$.

1.2. Tangent plane scheme (TPS). The idea of TPS [AJ06] can roughly be described as follows: At time t_n , the magnetization $\mathbf{m}(t_n)$ is discretized by some lowest-order finite element approximation \mathbf{m}_h^n in space. Discretizing (3) by a Galerkin approach in the discrete tangent space at \mathbf{m}_h^n , we obtain an approximation $\mathbf{v}_h^n \approx \mathbf{v}(t_n)$. Up to nodal normalization, $\mathbf{m}_h^n + k\mathbf{v}_h^n$ then yields an approximation of the magnetization at time $t_{n+1} := t_n + k$. Although LLG is nonlinear, TPS thus solves only one linear system per time-step for \mathbf{v}_h^n , yet, in the discrete tangent space.

1.3. State of the art. TPS with explicit time-stepping was first analyzed in [AJ06] with a refined analysis in [BKP08], which requires a CFL condition for convergence towards a weak solution in the sense of [AS92]. The work [Alo08] proposed TPS with an implicit time-stepping. This yields unconditional convergence of the algorithm towards a weak solution. While the algorithm of [Alo08] is formulated for the exchange field only, it was extended to general stationary lower-order contributions in [AKT12, BSF⁺14] and chiral magnetic skyrmion dynamics in [HPP⁺18]. Moreover, TPS was extended to the coupling of LLG with other evolution equations such as the full Maxwell system [BPP15], the eddy current equation [LT13, LPPT15, FT17], the conservation of elastic momentum [BPPR14] to model magnetostrictive effects, or a spin diffusion equation [AHP⁺14, ARB⁺15]. In the mentioned works, TPS is formally of first order in time. Recently, TPS was modified into a (formally) second-order in time scheme in [AKST14] with extensions in [DPP⁺17, HPP⁺18].

1.4. Contributions. So far, the efficient solution of the linear system in the discrete tangent space for the computation of $\mathbf{v}_h^n \approx \mathbf{v}(t_n) : \Omega \rightarrow \mathbb{R}^3$ has not been discussed in the literature, yet. Here, the main difficulty is the time-dependent ansatz space resulting in a time-dependent system matrix. This also aggravates the construction of suitable and effective preconditioners, which, if possible, should not depend on the time-step, or, at least, only need an update every once in a while (after several time-steps).

We construct a linear system in \mathbb{R}^{2N} , where $N \in \mathbb{N}$ is the number of nodes of the underlying finite element discretization. The corresponding system matrix is positive definite, but non-symmetric and depends on the time-step. We present and analyze various preconditioners, including a stationary approach (i.e., independent of the time-step) as well as Jacobi-type approximations. In the worst case, the number of necessary updates of the preconditioner to attain optimal convergence of the GMRES algorithm [SS86, Saa03] depends on the mesh-size h . However, under certain assumptions on the discrete magnetization $\mathbf{m}_h^n \approx \mathbf{m}(t_n)$, the number of necessary updates is also independent of h .

1.5. Outline. This paper is organized as follows: Section 2 introduces the basic notation and gives a precise formulation of TPS (Algorithm 1). In Section 3, we provide a basis for the discrete tangent space and derive the prototype linear system, which has to be solved in each time-step (Theorem 3). Section 4 proposes symmetric and positive definite preconditioners for the latter linear system. The two main results (Theorem 5 and Theorem 7) prove that the corresponding GMRES algorithms converge linearly. These results also provide estimates of the corresponding residual reduction factors (in certain energy norms) and show under which assumptions these estimates are independent of the discretization parameters. A corresponding linear convergence result for a stationary

preconditioning approach (Corollary 6) is a by-product of Theorem 5. Finally, we also discuss Jacobi-type approximations of our preconditioners (Section 4.4). Our theoretical results are underpinned by numerical experiments in Section 5. The proofs of Theorem 3, Theorem 5, and Theorem 7 are postponed to Section 6.

2. PRELIMINARIES

2.1. General notation. For any dimension $d \in \mathbb{N}$ (clear from the context) and vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, let $\mathbf{x} \cdot \mathbf{y}$ denote the Euclidean scalar product with the corresponding norm $|\mathbf{x}|^2 := \mathbf{x} \cdot \mathbf{x}$. The induced matrix norm reads $\|\mathbf{A}\| := \sup_{\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}} |\mathbf{A}\mathbf{x}|/|\mathbf{x}|$. Moreover, we denote by \mathbf{e}_i the i -th unit vector and by \mathbf{I} the identity matrix in \mathbb{R}^d . To abbreviate notation, we follow the `Matlab` syntax: For vectors, $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$, we write $[\mathbf{x}_1, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ for the matrix whose j -th column is \mathbf{x}_j . We use bold letters for vector-valued spaces, e.g., $\mathbf{L}^2(\Omega) = (L^2(\Omega))^3$. By slight abuse of notation, we write $\|\cdot\|_{\mathbf{L}^2(\Omega)}$ simultaneously for the L^2 -norm on $(L^2(\Omega))^3$ and $(L^2(\Omega))^{3 \times 3}$. Similarly, we write $\langle \cdot, \cdot \rangle_\Omega$ for all L^2 -scalar products including vector-valued spaces. We write ∂_k for the derivative with respect to \mathbf{x}_k , where $k \in \{1, 2, 3\}$. We write C for generic constants (clear from the context and, in particular, independent of the discretization parameters). For $a, b \in \mathbb{R}^+$ with $a \leq Cb$, we write $a \lesssim b$. If $a \lesssim b$ and $b \lesssim a$, we write $a \simeq b$.

2.2. Discretization. For the temporal discretization of LLG, let $M \in \mathbb{N}$ and $k := T/M$. Let $t_n := kn$ with $n \in \{0, \dots, M\}$ be the uniform time-steps and let $t_{n+1/2} := (t_{n+1} + t_n)/2$. For the spatial discretization, let \mathcal{T}_h be a C_{mesh} -quasi-uniform triangulation of Ω into tetrahedra $K \in \mathcal{T}_h$ with mesh-size $h > 0$, i.e., there exists $C_{\text{mesh}} > 0$ such that

$$h \leq |K|^{1/3} \leq \text{diam}(K) \leq C_{\text{mesh}} h \quad \text{for all } K \in \mathcal{T}_h,$$

where $|K|$ denotes the volume of the element K , while $\text{diam}(K)$ is its diameter. Let

$$\mathcal{S}_h := (\mathcal{S}_h)^3 \quad \text{with} \quad \mathcal{S}_h := \{v_h \in C(\Omega) : \text{for all } T \in \mathcal{T}_h \quad v_h|_T \text{ is affine}\} \quad (4)$$

be the lowest-order FEM space. We denote by \mathcal{N}_h the set of nodes of \mathcal{T}_h and define $N := \#\mathcal{N}_h$. For all $d \in \mathbb{N}$, a scaling argument (see, e.g., [Bar15, Lemma 3.9]) yields that

$$\|\varphi_h\|_{\mathbf{L}^2(\Omega)}^2 \leq h^3 \sum_{i=1}^N |\varphi_h(z_i)|^2 \leq (d+2) \|\varphi_h\|_{L^2(\Omega)}^2 \quad \text{for all } \varphi_h \in (\mathcal{S}_h)^d. \quad (5)$$

Moreover, let

$$\mathcal{M}_h := \{\varphi_h \in \mathcal{S}_h : |\varphi_h(z)| = 1 \quad \text{for all } z \in \mathcal{N}_h\}. \quad (6)$$

For some fixed $\boldsymbol{\mu}_h \in \mathcal{M}_h$, define the discrete tangent space

$$\mathcal{K}_h[\boldsymbol{\mu}_h] := \{\varphi_h \in \mathcal{S}_h : \varphi_h(z) \cdot \boldsymbol{\mu}_h(z) = 0 \quad \text{for all } z \in \mathcal{N}_h\}. \quad (7)$$

Note that $\dim \mathcal{S}_h = 3N$ and $\dim \mathcal{K}_h[\boldsymbol{\mu}_h] = 2N$.

2.3. Tangent plane scheme (TPS). We unify the formulation of the first-order TPS [Alo08, AKT12, BSF⁺14] and the second-order TPS [AKST14, DPP⁺17]. To this end, let $W_k : \mathbb{R} \rightarrow \mathbb{R}_{>0}$ and $\beta : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ satisfy

$$\lim_{k \rightarrow 0} \|W_k - \alpha\|_{L^\infty(\mathbb{R})} = 0, \quad \inf_{k > 0} \inf_{s \in \mathbb{R}} W_k(s) \geq \frac{\alpha}{2} > 0, \quad \text{and} \quad \lim_{k \rightarrow 0} \beta(k)k = 0. \quad (8)$$

Then, the general algorithm takes the following form:

Algorithm 1 (General TPS). **Input:** $\mathbf{m}_h^{-1} := \mathbf{m}_h^0 \in \mathcal{M}_h$.

Loop: For $0 \leq n < M$, iterate the following steps (a)–(c):

(a) Compute

$$\lambda_h^n := -\ell_{\text{ex}}^2 |\nabla \mathbf{m}_h^n|^2 + (\mathbf{f}(t_n) + \boldsymbol{\pi}(\mathbf{m}_h^n)) \cdot \mathbf{m}_h^n. \quad (9a)$$

(b) Find $\mathbf{v}_h^n \in \mathcal{K}_h[\mathbf{m}_h^n]$ such that

$$\begin{aligned} & \langle W_k(\lambda_h^n) \mathbf{v}_h^n, \boldsymbol{\varphi}_h \rangle_\Omega + \langle \mathbf{m}_h^n \times \mathbf{v}_h^n, \boldsymbol{\varphi}_h \rangle_\Omega + \beta(k)k \langle \nabla \mathbf{v}_h^n, \nabla \boldsymbol{\varphi}_h \rangle_\Omega \\ & = -\ell_{\text{ex}}^2 \langle \nabla \mathbf{m}_h^n, \nabla \boldsymbol{\varphi}_h \rangle_\Omega + \langle \mathbf{L}_h(\mathbf{m}_h^n, \mathbf{m}_h^{n-1}), \boldsymbol{\varphi}_h \rangle_\Omega \quad \text{for all } \boldsymbol{\varphi}_h \in \mathcal{K}_h[\mathbf{m}_h^n]; \end{aligned} \quad (9b)$$

see Remark 2 (iii) for details on W_k , β , and $\mathbf{L}_h(\mathbf{m}_h^n, \mathbf{m}_h^{n-1}) \in \mathbf{H}^1(\Omega)$.

(c) Define $\mathbf{m}_h^{n+1} \in \mathcal{M}_h$ by

$$\mathbf{m}_h^{n+1}(z) := \frac{\mathbf{m}_h^n(z) + k\mathbf{v}_h^n(z)}{|\mathbf{m}_h^n(z) + k\mathbf{v}_h^n(z)|} \quad \text{for all nodes } z \in \mathcal{N}_h. \quad (9c)$$

Output: Approximations $\mathbf{m}_h^n \approx \mathbf{m}(t_n)$. □

Remark 2. (i) With $W_k(\cdot) \geq \alpha/2 > 0$ from (8), the bilinear form on the left-hand side of (9b) is continuous and elliptic on $\mathbf{H}^1(\Omega)$. Therefore, the Lax–Milgram theorem guarantees existence and uniqueness of the solution $\mathbf{v}_h^n \in \mathcal{K}_h[\mathbf{m}_h^n]$ to (9b).

(ii) To see that (9c) is well-defined, note that $\mathbf{m}_h^0 \in \mathcal{M}_h$ and induction on n prove that

$$|\mathbf{m}_h^n(z) + k\mathbf{v}_h^n(z)|^2 \stackrel{(7)}{=} |\mathbf{m}_h^n(z)|^2 + k^2 |\mathbf{v}_h^n(z)|^2 \geq 1 \quad \text{for all nodes } z \in \mathcal{N}_h.$$

(iii) The first-order (TPS1) and second-order (TPS2) tangent plane schemes differ solely in W_k , β and $\mathbf{L}_h(\mathbf{m}_h^n, \mathbf{m}_h^{n-1})$ in the definition of (9b). For TPS1 from [Alo08, AKT12, BSF⁺14], we have

$$W_k \equiv \alpha, \quad (10a)$$

$$\beta(k) \equiv \ell_{\text{ex}}^2 \Theta \quad \text{with } \Theta \in (0, 1], \quad (10b)$$

$$\mathbf{L}_h(\mathbf{m}_h^n, \mathbf{m}_h^{n-1}) := \boldsymbol{\pi}(\mathbf{m}_h^n) + \mathbf{f}(t_n), \quad (10c)$$

and step (a) in Algorithm 1 is omitted. For TPS2 from [AKST14, DPP⁺17], let

$$\rho(k) := |k \log(k)| \quad \text{and} \quad M(k) := |k \log(k)|^{-1}; \quad (11a)$$

see, e.g., [DPP⁺17, eq.(11c)] for the precise requirements. Then, we have

$$W_k(s) := \begin{cases} \alpha + \frac{k}{2} \min\{s, M(k)\} & \text{for } s \geq 0, \\ \frac{\alpha}{1 + \frac{k}{2\alpha} \min\{-s, M(k)\}} & \text{for } s < 0, \end{cases} \quad (11b)$$

$$\beta(k) := \frac{\ell_{\text{ex}}^2}{2} (1 + \rho(k)), \quad (11c)$$

$$\mathbf{L}_h(\mathbf{m}_h^n, \mathbf{m}_h^{n-1}) := \frac{3}{2} \boldsymbol{\pi}(\mathbf{m}_h^n) - \frac{1}{2} \boldsymbol{\pi}(\mathbf{m}_h^{n-1}) + \mathbf{f}\left(t_n + \frac{k}{2}\right). \quad (11d)$$

Note that, for sufficiently small $k > 0$, the latter choices of W_k and ρ satisfy the assumptions from (8); see [AKST14, DPP⁺17] for details.

(iv) For TPS1 from (iii), the projection step (c) of Algorithm 1 can be omitted. The integrator then remains unconditionally convergent [AHP⁺14]. If there exists a smooth (and hence unique [DS14]) strong solution of LLG, the a priori analysis in [FT17] guarantees first-order convergence in space and time.

2.4. Linear algebra. We suppose a numbering of the nodes, i.e., $\mathcal{N}_h = \{z_1, \dots, z_N\}$. Let $\varphi_j \in \mathcal{S}_h$ be the nodal hat function associated with z_j , i.e., $\varphi_j(z_k) = \delta_{jk}$, where δ_{jk} is Kronecker's delta. We then consider the following basis of \mathcal{S}_h : Define

$$\phi_{3(j-1)+\ell} := \varphi_j \mathbf{e}_\ell \quad \text{for all } j = 1, \dots, N \text{ and all } \ell = 1, 2, 3. \quad (12)$$

Given $\mathbf{m}_h^n \in \mathcal{M}_h$, we then define $\mathbf{M}, \mathbf{M}_k[\mathbf{m}_h^n], \mathbf{L}, \mathbf{S}[\mathbf{m}_h^n] \in \mathbb{R}^{3N \times 3N}$ as follows:

- $\mathbf{M}_{ij} := \langle \phi_i, \phi_j \rangle_\Omega$ is the (symmetric and positive definite) mass matrix;
- $(\mathbf{M}_k[\mathbf{m}_h^n])_{ij} := \langle (W_k(\lambda_h^n)/\alpha) \phi_i, \phi_j \rangle_\Omega$ is the (symmetric and positive definite) weighted mass matrix, where λ_h^n stems from (9a) and thus depends on \mathbf{m}_h^n ;
- $\mathbf{L}_{ij} := \langle \nabla \phi_i, \nabla \phi_j \rangle_\Omega$ is the (symmetric and positive semidefinite) stiffness matrix;
- $(\mathbf{S}[\mathbf{m}_h^n])_{ij} := \langle \mathbf{m}_h^n \times \phi_i, \phi_j \rangle_\Omega$ is the (skew-symmetric) cross product matrix.

Moreover, we set

$$M_{ij} := \langle \varphi_i, \varphi_j \rangle_\Omega \in \mathbb{R} \quad \text{and} \quad L_{ij} := \langle \nabla \varphi_i, \nabla \varphi_j \rangle_\Omega \in \mathbb{R} \quad \text{for } i, j = 1, \dots, N, \quad (13a)$$

and note the block forms

$$\mathbf{M} = \begin{pmatrix} M_{11} \mathbf{I}_{3 \times 3} & \cdots & M_{1N} \mathbf{I}_{3 \times 3} \\ \vdots & \ddots & \vdots \\ M_{N1} \mathbf{I}_{3 \times 3} & \cdots & M_{NN} \mathbf{I}_{3 \times 3} \end{pmatrix} \quad \text{and} \quad \mathbf{L} = \begin{pmatrix} L_{11} \mathbf{I}_{3 \times 3} & \cdots & L_{1N} \mathbf{I}_{3 \times 3} \\ \vdots & \ddots & \vdots \\ L_{N1} \mathbf{I}_{3 \times 3} & \cdots & L_{NN} \mathbf{I}_{3 \times 3} \end{pmatrix}. \quad (13b)$$

If we replace $\mathcal{K}_h[\mathbf{m}_h^n]$ with \mathcal{S}_h in (9b), the left-hand side of (9b) gives rise to the matrix

$$\mathbf{A}_k[\mathbf{m}_h^n] := \alpha \mathbf{M}_k[\mathbf{m}_h^n] + \beta(k)k \mathbf{L} - \mathbf{S}[\mathbf{m}_h^n] \in \mathbb{R}^{3N \times 3N}. \quad (14)$$

The right-hand side of (9b) gives rise to the vector $\mathbf{b}[\mathbf{m}_h^n] \in \mathbb{R}^{3N}$ with

$$(\mathbf{b}[\mathbf{m}_h^n])_j := -\ell_{\text{ex}}^2 \langle \nabla \mathbf{m}_h^n, \nabla \phi_j \rangle_\Omega + \langle \mathbf{L}_h(\mathbf{m}_h^n, \mathbf{m}_h^{n-1}), \varphi_h \rangle_\Omega. \quad (15)$$

Note that the matrix $\mathbf{A}_k[\mathbf{m}_h^n]$ is positive definite and hence regular, but not symmetric. Finally, define the 2D-equivalent to the basis from (12) by

$$\psi_{2(j-1)+\ell} := \varphi_j \mathbf{e}_\ell \quad \text{for all } j = 1, \dots, N \text{ and all } \ell = 1, 2. \quad (16)$$

3. THE TANGENT SPACE PROBLEM

In this section, we present a strategy, which translates the solution of the discrete variational formulation (9b) to a linear system in $\mathbb{R}^{2N} \cong \mathcal{K}_h[\mathbf{m}_h^n]$. To that end, we use Householder matrices: Given $\mathbf{m} \in \mathbb{R}^3$ with $|\mathbf{m}| = 1$, define $\tilde{\mathbf{H}}[\mathbf{m}] \in \mathbb{R}^{3 \times 3}$ by

$$\tilde{\mathbf{H}}[\mathbf{m}] := \begin{cases} \mathbf{I} - 2\mathbf{w}\mathbf{w}^T, & \text{where } \mathbf{w} := \frac{\mathbf{m} + \mathbf{e}_3}{|\mathbf{m} + \mathbf{e}_3|} \quad \text{for } \mathbf{m} \neq -\mathbf{e}_3, \\ [\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_3] & \text{for } \mathbf{m} = -\mathbf{e}_3. \end{cases} \quad (17a)$$

Then, $\tilde{\mathbf{H}}[\mathbf{m}]$ is orthonormal with $\tilde{\mathbf{H}}[\mathbf{m}] = \tilde{\mathbf{H}}[\mathbf{m}]^T = \tilde{\mathbf{H}}[\mathbf{m}]^{-1}$ and maps \mathbf{e}_3 to $-\mathbf{m}$. Define

$$\mathbf{H}[\mathbf{m}] := [\tilde{\mathbf{H}}[\mathbf{m}]\mathbf{e}_1, \tilde{\mathbf{H}}[\mathbf{m}]\mathbf{e}_2] \in \mathbb{R}^{3 \times 2}, \quad (17b)$$

i.e., $\text{span}(\mathbf{H}[\mathbf{m}]) \perp \mathbf{m}$ and $\text{span}(\mathbf{H}[\mathbf{m}])$ mimics $\mathcal{K}_h[\mathbf{m}_h^n]$ nodewise. Moreover, for any orthogonal matrix $\mathbf{T}_n \in \mathbb{R}^{3 \times 3}$ with $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$, the matrix

$$\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \mathbf{m}] \in \mathbb{R}^{3 \times 2} \quad \text{instead of} \quad \mathbf{H}[\mathbf{m}] \in \mathbb{R}^{3 \times 2}, \quad (17c)$$

also satisfies $\text{span}(\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \mathbf{m}]) \perp \mathbf{m}$. Hence, $\text{span}(\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \mathbf{m}])$ still mimics $\mathcal{K}_h[\mathbf{m}_h^n]$ nodewise. The following theorem provides a linear system in $\mathbb{R}^{2N} \cong \mathcal{K}_h[\mathbf{m}_h^n]$ for the solution to (9b). The proof is postponed to Section 6.3 below.

Theorem 3. Recall $\mathbf{A}_k[\mathbf{m}_h^n] \in \mathbb{R}^{3N \times 3N}$ and $\mathbf{b}[\mathbf{m}_h^n] \in \mathbb{R}^{3N}$ from Section 2.4. Define

$$\mathbf{Q}[\mathbf{m}_h^n] := \begin{pmatrix} \mathbf{T}_n \mathbf{H}[\mathbf{T}_n \mathbf{m}_h^n(z_1)] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_n \mathbf{H}[\mathbf{T}_n \mathbf{m}_h^n(z_2)] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{T}_n \mathbf{H}[\mathbf{T}_n \mathbf{m}_h^n(z_N)] \end{pmatrix} \in \mathbb{R}^{3N \times 2N}, \quad (18)$$

where $\mathbf{H}(\cdot)$ stems from (17). Then, the matrix $\mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{A}_k[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n] \in \mathbb{R}^{2N \times 2N}$ is positive definite and, in particular, regular. Moreover, the unique solution $\mathbf{x} \in \mathbb{R}^{2N}$ of

$$(\mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{A}_k[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n]) \mathbf{x} = \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{b}[\mathbf{m}_h^n], \quad (19)$$

and the unique solution $\mathbf{v}_h^n \in \mathcal{K}_h[\mathbf{m}_h^n]$ of the variational formulation (9b) satisfy

$$\mathbf{v}_h^n = \sum_{j=1}^{3N} \mathbf{v}_j \phi_j \quad \text{with} \quad \mathbf{v} := \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \in \mathbb{R}^{3N}. \quad (20)$$

Remark 4. For the validity of Theorem 3, it is only relevant that $\mathbf{H}[\mathbf{m}_h^n(z_i)]$ has orthonormal columns and that $\text{span}(\mathbf{H}[\mathbf{m}_h^n(z_i)]) \perp \mathbf{m}_h^n(z_i)$ for all $i = 1, \dots, N$. Given $\mathbf{m} \in \mathbb{R}^3$, with $|\mathbf{m}| = 1$, alternative strategies from [Rug16, Lemma 6.1.2] are, e.g.,

- either to set $\mathbf{w} := \frac{\mathbf{m} + \sigma \mathbf{e}_3}{|\mathbf{m} + \sigma \mathbf{e}_3|}$, where $\sigma = \text{sign}(\mathbf{m}_3)$;
- or to use the transformation matrix of the rotation around the axis $\mathbf{e}_3 \times \mathbf{m}$ by an angle φ such that $\cos \varphi = \mathbf{m} \cdot \mathbf{e}_3$ and $\sin \varphi = |\mathbf{m} \times \mathbf{e}_3|$.

Moreover, even different strategies for $\mathbf{H}[\mathbf{m}_h^n(z_i)]$ at different nodes z_i are possible.

4. PRECONDITIONING

To solve the tangent space system (19), we aim to choose a preconditioner $\mathbf{P} \in \mathbb{R}^{2N \times 2N}$ and employ the GMRES algorithm [SS86, Saa03] to the preconditioned system

$$\mathbf{P} \mathbf{A}_{\mathbf{Q}}[\mathbf{m}_h^n] \mathbf{x} := \mathbf{P} \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{A}_k[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} = \mathbf{P} \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{b}[\mathbf{m}_h^n] =: \mathbf{P} \mathbf{b}_{\mathbf{Q}}[\mathbf{m}_h^n]. \quad (21)$$

In the following sections, we discuss possible choices for \mathbf{P} . We rely on the symmetric part of $\mathbf{A}_{\mathbf{Q}}[\mathbf{m}_h^n]$, where we replace W_k by the parameter

$$\alpha_{\mathbf{P}} > 0. \quad (22)$$

In particular, this includes the case $\alpha_{\mathbf{P}} = \alpha$. Note that GMRES requires only the action of the preconditioner \mathbf{P} on a vector. Moreover, recall that $\mathbf{Q}[\mathbf{m}_h^n]$ from (18) implicitly depends on the arbitrary but fixed matrix $\mathbf{T}_n \in \mathbb{R}^{3 \times 3}$ from (17c). We refer to Section 4.5 below for the possible construction of the matrix \mathbf{T}_n , for given $\mathbf{m}_h^n \in \mathcal{M}_h$.

4.1. Theoretical preconditioner. For $\boldsymbol{\mu}_h \in \mathcal{M}_h$, we first consider

$$\mathbf{P}_{\mathbf{Q}}[\boldsymbol{\mu}_h] := \left(\mathbf{Q}[\boldsymbol{\mu}_h]^T (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k \mathbf{L}) \mathbf{Q}[\boldsymbol{\mu}_h] \right)^{-1} \in \mathbb{R}^{2N \times 2N}. \quad (23)$$

To analyse the preconditioned GMRES algorithm, we define the energy scalar product

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle_{\boldsymbol{\mu}_h} := \mathbf{x} \cdot \left(\mathbf{P}_{\mathbf{Q}}[\boldsymbol{\mu}_h] \right)^{-1} \mathbf{y} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N} \quad (24)$$

and denote the induced norm by $\|\cdot\|_{\boldsymbol{\mu}_h}$. The following theorem shows that (the time-step dependent) $\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$ is a suitable choice and that, in practice, one can keep and reuse $\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$ for several time-steps. Its proof is postponed to Section 6.4.

Theorem 5. *Let $\alpha_{\mathbf{P}} \geq \alpha$. Let $\boldsymbol{\mu}_h \in \mathcal{M}_h$ be arbitrary. Consider the preconditioned GMRES algorithm with the preconditioner $\mathbf{P}_{\mathbf{Q}}[\boldsymbol{\mu}_h]$ from (23) for the solution of (21) with the initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^{2N}$. For $\ell \in \mathbb{N}_0$, let $\mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}$ denote the GMRES iterates with the corresponding residuals*

$$\mathbf{r}^{(\ell)} := \mathbf{P}_{\mathbf{Q}}[\boldsymbol{\mu}_h] \mathbf{b}_{\mathbf{Q}}[\mathbf{m}_h^n] - \mathbf{P}_{\mathbf{Q}}[\boldsymbol{\mu}_h] \mathbf{A}_{\mathbf{Q}}[\mathbf{m}_h^n] \mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}.$$

Then, the following two assertions (i)–(ii) hold true:

(i) *There exists a constant $C > 1$, which depends only on C_{mesh} , such that*

$$\|\mathbf{r}^{(\ell)}\|_{\boldsymbol{\mu}_h} \leq \left[1 - C^{-1} \left(\frac{\alpha}{\alpha_{\mathbf{P}}} \right)^2 \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right)^{-2} \mathbb{F}^{-4} \right]^{\ell/2} \|\mathbf{r}^{(0)}\|_{\boldsymbol{\mu}_h} \quad (25)$$

for all $\ell \in \mathbb{N}_0$, where

$$\mathbb{F} := 1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} h^2} \max_{z \in \mathcal{N}_h} \|\mathbf{H}[\mathbf{T}_n \mathbf{m}_h^n(z)] - \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z)]\|^2 \geq 1. \quad (26)$$

(ii) *If, additionally, $1 + (\mathbf{T}_n \mathbf{m}_h^n(z))_3 \geq \gamma > 0$ and $1 + (\mathbf{T}_n \boldsymbol{\mu}_h(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, the statement of (i) holds with the h -independent factor*

$$\begin{aligned} \mathbb{F} &:= 1 + \gamma^{-2} \|\mathbf{m}_h^n - \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}^2 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^2} \|\nabla \mathbf{m}_h^n - \nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}^2 \\ &+ \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^6} \left(\|\nabla \mathbf{m}_h^n\|_{\mathbf{L}^\infty(\Omega)}^2 + \|\nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}^2 \right) \|\mathbf{m}_h^n - \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}^2 \geq 1. \end{aligned} \quad (27)$$

4.2. Stationary preconditioning. We consider a preconditioner which is independent of the time-step. Similarly to $\mathbf{M}, \mathbf{L} \in \mathbb{R}^{3N \times 3N}$ from Section 2.4, define the matrices $\mathbf{M}^{2D}, \mathbf{L}^{2D} \in \mathbb{R}^{2N \times 2N}$, which correspond to the nodal basis $(\boldsymbol{\psi}_i)_{i=1}^{2N}$ of $(\mathcal{S}_h)^2$ from (16):

- $\mathbf{M}_{ij}^{2D} := \langle \boldsymbol{\psi}_i, \boldsymbol{\psi}_j \rangle_{\Omega}$ is the (symmetric and positive definite) mass matrix,
- $\mathbf{L}_{ij}^{2D} := \langle \nabla \boldsymbol{\psi}_i, \nabla \boldsymbol{\psi}_j \rangle_{\Omega}$ is the (symmetric and positive semidefinite) stiffness matrix.

Then, consider the stationary preconditioner

$$\mathbf{P}^{2D} := \left(\alpha_{\mathbf{P}} \mathbf{M}^{2D} + \beta(k)k \mathbf{L}^{2D} \right)^{-1} \in \mathbb{R}^{2N \times 2N}. \quad (28)$$

Denote the corresponding energy scalar product by

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle := \mathbf{x} \cdot (\alpha_{\mathbf{P}} \mathbf{M}^{2\text{D}} + \beta(k)k\mathbf{L}^{2\text{D}}) \mathbf{y} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N},$$

and the induced norm by $\|\cdot\|$. The following corollary discusses the performance of $\mathbf{P}^{2\text{D}}$. It is a direct consequence of Theorem 5(ii).

Corollary 6. *Let $\alpha_{\mathbf{P}} \geq \alpha$. Consider the preconditioned GMRES algorithm with $\mathbf{P}^{2\text{D}}$ from (28) for the solution of (21) with the initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^{2N}$. For $\ell \in \mathbb{N}_0$, let $\mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}$ denote the GMRES iterates with the corresponding residuals*

$$\mathbf{r}^{(\ell)} := \mathbf{P}^{2\text{D}} \mathbf{b}_{\mathbf{Q}}[\mathbf{m}_h^n] - \mathbf{P}^{2\text{D}} \mathbf{A}_{\mathbf{Q}}[\mathbf{m}_h^n] \mathbf{x}^{(\ell)} \in \mathbb{R}^{2N}.$$

Let $1 + (\mathbf{T}_n \mathbf{m}_h^n(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. Then, there exists a constant $C > 1$, which depends only on C_{mesh} , such that

$$\|\mathbf{r}^{(\ell)}\| \leq \left[1 - C^{-1} \left(\frac{\alpha}{\alpha_{\mathbf{P}}} \right)^2 \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right)^{-2} \mathbb{F}^{-4} \right]^{\ell/2} \|\mathbf{r}^{(0)}\| \quad (29)$$

for all $\ell \in \mathbb{N}_0$, where

$$\mathbb{F} := 1 + \gamma^{-2} + \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^6} \|\nabla \mathbf{m}_h^n\|_{\mathbf{L}^\infty(\Omega)}^2 \geq 1. \quad (30)$$

Proof. Recall that $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$. For constant $\boldsymbol{\mu}_h := \mathbf{T}_n \mathbf{e}_3 \in \mathcal{M}_h$, we get that

$$\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_i)] = \mathbf{T}_n \mathbf{H}[\mathbf{e}_3] = \mathbf{T}_n[\mathbf{e}_1, \mathbf{e}_2] \in \mathbb{R}^{3 \times 2} \quad (31a)$$

and thus

$$(\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_i)])^T (\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_j)]) = [\mathbf{e}_1, \mathbf{e}_2]^T \mathbf{T}_n \mathbf{T}_n [\mathbf{e}_1, \mathbf{e}_2] = \mathbf{I}_{2 \times 2}. \quad (31b)$$

for all $i, j = 1, \dots, N$. Together with the block forms from (13), this yields that

$$\mathbf{P}_{\mathbf{Q}}[\boldsymbol{\mu}_h] \stackrel{(23)}{=} (\mathbf{Q}[\boldsymbol{\mu}_h]^T (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{Q}[\boldsymbol{\mu}_h])^{-1} \stackrel{(31)}{=} (\alpha_{\mathbf{P}} \mathbf{M}^{2\text{D}} + \beta(k)k\mathbf{L}^{2\text{D}})^{-1} \stackrel{(28)}{=} \mathbf{P}^{2\text{D}}.$$

Since $\nabla \boldsymbol{\mu}_h = \mathbf{0}$ a.e. in Ω and $\|\mathbf{m}_h^n - \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} \leq 2$, Theorem 5(ii) proves the result. \square

4.3. Practical preconditioner. For general problems of type (19), the work [NS96] proposes (without a proof) to consider the practical preconditioner

$$\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n] := \mathbf{Q}[\mathbf{m}_h^n]^T (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L})^{-1} \mathbf{Q}[\mathbf{m}_h^n] \in \mathbb{R}^{2N \times 2N} \quad (32)$$

to approximate the theoretical preconditioner $\mathbf{P}[\mathbf{m}_h^n]$ from (23). We note that unlike $\mathbf{P}_{\mathbf{Q}}[\boldsymbol{\mu}_h]$, it makes no sense to consider $\tilde{\mathbf{P}}_{\mathbf{Q}}[\boldsymbol{\mu}_h]$ for $\boldsymbol{\mu}_h \neq \mathbf{m}_h^n$, since the preconditioned GMRES algorithm exploits only the matrix-vector product with the preconditioner and $\mathbf{Q}[\mathbf{m}_h^n]$ is computed anyway. For the analysis, we define the energy scalar product

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle'_{\mathbf{m}_h^n} := \mathbf{x} \cdot (\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n])^{-1} \mathbf{y} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}, \quad (33)$$

and denote the corresponding norm by $\|\cdot\|'_{\mathbf{m}_h^n}$. The following theorem discusses the performance of preconditioned GMRES with the practical preconditioner $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$. Its proof is postponed to Section 6.5.

Theorem 7. Let $\alpha_{\mathbf{P}} \geq \alpha$. Consider the preconditioned GMRES algorithm with the preconditioner $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ from (32) for the solution of (21) with the initial guess $\tilde{\mathbf{x}}^{(0)} \in \mathbb{R}^{2N}$. For $\ell \in \mathbb{N}_0$, let $\tilde{\mathbf{x}}^{(\ell)} \in \mathbb{R}^{2N}$ denote the GMRES iterates with the corresponding residuals

$$\tilde{\mathbf{r}}^{(\ell)} := \tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n] \mathbf{b}_{\mathbf{Q}}[\mathbf{m}_h^n] - \tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n] \mathbf{A}_{\mathbf{Q}}[\mathbf{m}_h^n] \tilde{\mathbf{x}}^{(\ell)} \in \mathbb{R}^{2N}.$$

Then, the following two assertions (i)–(ii) hold true:

(i) There exists a constant $C > 1$, which depends only on C_{mesh} , such that

$$\|\tilde{\mathbf{r}}^{(\ell)}\|'_{\mathbf{m}_h^n} \leq \left[1 - C^{-1} \left(\frac{\alpha}{\alpha_{\mathbf{P}}} \right)^2 \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right)^{-2} \tilde{\mathbb{F}}^{-1} \right]^{\ell/2} \|\tilde{\mathbf{r}}^{(0)}\|'_{\mathbf{m}_h^n} \quad (34)$$

for all $\ell \in \mathbb{N}_0$, where

$$\tilde{\mathbb{F}} := 1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}}h^2} \geq 1. \quad (35)$$

(ii) If, additionally, $1 + (\mathbf{T}_n \mathbf{m}_h^n(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, the statement of (i) holds with the h -independent factor

$$\tilde{\mathbb{F}} := 1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}}\gamma^4} \|\nabla \mathbf{m}_h^n\|_{L^\infty(\Omega)}^2 \geq 1. \quad (36)$$

4.4. Jacobi-type preconditioner. Consider the following approximation to the stationary preconditioner \mathbf{P}^{2D} from (28): Recalling $M_{ij}, L_{ij} \in \mathbb{R}$ from (13a), we set

$$\mathbf{P}_i^{\text{jac}} := (\alpha_{\mathbf{P}} M_{ii} + \beta(k)k L_{ii})^{-1} \mathbf{I}_{2 \times 2} \in \mathbb{R}^{2 \times 2} \quad \text{for all } i = 1, \dots, N \quad (37)$$

and define the stationary Jacobi-type preconditioner

$$\mathbf{P}^{\text{jac}} := \begin{pmatrix} \mathbf{P}_1^{\text{jac}} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2^{\text{jac}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{P}_N^{\text{jac}} \end{pmatrix} \in \mathbb{R}^{2N \times 2N}. \quad (38)$$

Given $\mathbf{m}_h^n \in \mathcal{M}_h$, the matrix $\mathbf{P}_{\mathbf{Q}}^{\text{jac}}[\mathbf{m}_h^n] \in \mathbb{R}^{2N \times 2N}$

$$\left[(\mathbf{P}_{\mathbf{Q}}^{\text{jac}}[\mathbf{m}_h^n])^{-1} \right]_{ij} := \begin{cases} (\mathbf{Q}[\mathbf{m}_h^n]^T (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k \mathbf{L}) \mathbf{Q}[\mathbf{m}_h^n])_{ii} & \text{for } i = j, \\ 0 & \text{else,} \end{cases}$$

is the Jacobi-type approximation of $\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$ from (23). Similarly, the matrix

$$\mathbf{Q}[\mathbf{m}_h^n]^T \tilde{\mathbf{P}}^{\text{jac}} \mathbf{Q}[\mathbf{m}_h^n] \in \mathbb{R}^{2N \times 2N}, \quad \text{where } \left[(\tilde{\mathbf{P}}^{\text{jac}})^{-1} \right]_{ij} := \begin{cases} (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k \mathbf{L})_{ii} & \text{for } i = j, \\ 0 & \text{else,} \end{cases}$$

is the Jacobi-type approximation of $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ from (32). The following proposition states that all three definitions of Jacobi-type preconditioners coincide.

Proposition 8. It holds that

$$\mathbf{P}^{\text{jac}} = \mathbf{P}_{\mathbf{Q}}^{\text{jac}}[\mathbf{m}_h^n] = \mathbf{Q}[\mathbf{m}_h^n]^T \tilde{\mathbf{P}}^{\text{jac}} \mathbf{Q}[\mathbf{m}_h^n].$$

Proof. The statement follows from the representations of \mathbf{M} and \mathbf{L} from (13) as well as the block-diagonal definition of $\mathbf{Q}[\mathbf{m}_h^n]$, where the blocks $\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \mathbf{m}_h^n(z_i)]$ have orthonormal columns for all nodes $z_i \in \mathcal{N}_h$. \square

4.5. Practical computation of \mathbf{T}_n . With $\gamma > 0$ being as large as possible, Theorem 5 (ii), Corollary 6, and Theorem 7 (ii) require that

$$1 + \mathbf{m}_h^n(z) \cdot \mathbf{T}_n \mathbf{e}_3 \geq \gamma > 0 \quad \text{for all nodes } z \in \mathcal{N}_h. \quad (39)$$

For given $\mathbf{m}_h^n \in \mathcal{M}_h$, we thus aim to choose the matrix $\mathbf{T}_n \in \mathbb{R}^{3 \times 3}$ from (17c) such that γ is large. To this end, define $d_\ell^* \in [0, 2]$ by

$$d_\ell^+ := 1 - \max_{z \in \mathcal{N}_h} (\mathbf{m}_h^n(z))_\ell, \quad d_\ell^- := 1 + \min_{z \in \mathcal{N}_h} (\mathbf{m}_h^n(z))_\ell, \quad \text{for all } \ell \in \{1, 2, 3\}. \quad (40a)$$

In addition, define $\mathbf{T}_\ell^* \in \mathbb{R}^{3 \times 3}$ with $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$ by

$$\mathbf{T}_1^+ := [-\mathbf{e}_3, \mathbf{e}_2, -\mathbf{e}_1], \quad \mathbf{T}_1^- := [\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1], \quad (41a)$$

$$\mathbf{T}_2^+ := [\mathbf{e}_1, -\mathbf{e}_3, -\mathbf{e}_2], \quad \mathbf{T}_2^- := [\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2], \quad (41b)$$

$$\mathbf{T}_3^+ := [\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_3], \quad \mathbf{T}_3^- := [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]. \quad (41c)$$

For all $\ell \in \{1, 2, 3\}$ and $\star \in \{+, -\}$, it holds that

$$1 + \mathbf{m}_h^n(z) \cdot \mathbf{T}_\ell^* \mathbf{e}_3 \geq d_\ell^* \in [0, 2] \quad \text{for all nodes } z \in \mathcal{N}_h. \quad (42)$$

Hence, (39) holds with $\gamma \in [0, 2]$ being the maximum d_ℓ^* , and $\mathbf{T}_n = \mathbf{T}_\ell^*$ in (41). We note that other (more involved) strategies are possible. Finally, we note that our construction leads to $\gamma = 0$, if and only if

$$\{\pm \mathbf{e}_1, \pm \mathbf{e}_2, \pm \mathbf{e}_3\} \subset \{\mathbf{m}_h^n(z) : z \in \mathcal{N}_h\}.$$

5. NUMERICS

In this section, we underpin our theoretical findings with numerical experiments. To this end, we employ our C++-code for computational micromagnetics, which is based on the FEM library `NGSolve` [ngs]. For BEM computations, we employ the `BEM++` library [SBA⁺15]. Moreover, we couple `NGSolve` and `BEM++` via `ngbem` [Rie].

We always employ the first-order tangent plane scheme, i.e., Algorithm 1 with W_k , β , and $\mathbf{L}_h(\mathbf{m}_h^n, \mathbf{m}_h^{n-1})$ as in (10), where, in particular, we always set $\Theta = 1$ in (10b).

To solve the (preconditioned) linear system (19), we employ the (preconditioned) GMRES algorithm [SS86, Saa03]. Our implementation is based on the template routine from `Netlib` [gmr], where we employ the iteration tolerance $\varepsilon = 10^{-14}$. To save memory, we restart GMRES after every 200 iterations. Note that this is commonly referred to as restarted GMRES; cf., e.g., [Saa03, Algorithm 6.11]. As initial value for the GMRES iteration, we always choose $\mathbf{x}_0 = \mathbf{0}$.

The experiments of this section focus on the (possibly h -independent) iteration numbers of the preconditioned GMRES algorithm. To evaluate the inverse matrices in the preconditioners from Section 4 (e.g., in (23)), we always solve the corresponding linear system via Gaussian elimination.

Remark 9. *The corresponding 'inverse' matrices in the stationary preconditioner \mathbf{P}^{2D} from (28) as well as the practical preconditioner $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ from (32) are block-diagonal*

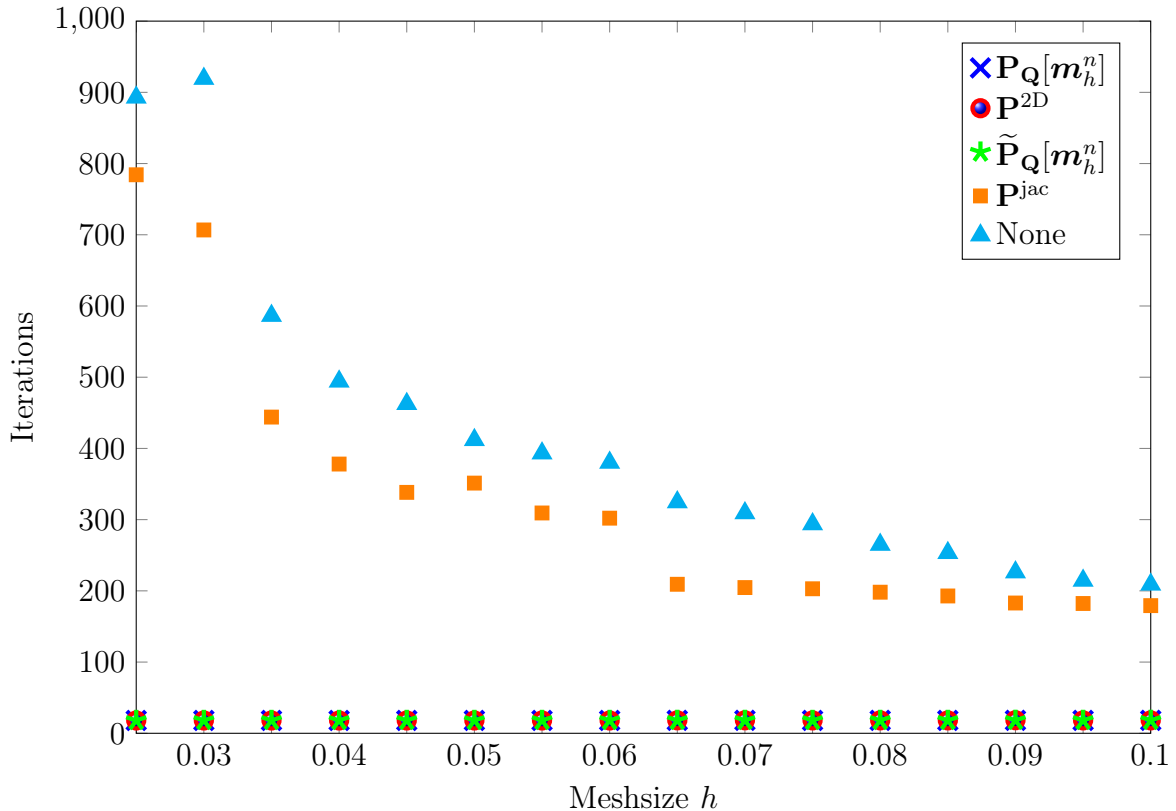


FIGURE 1. Experiment of Section 5.1: Average number of GMRES iterations.

matrices and consist of the same $N \times N$ -stationary matrix block. Moreover, $\mathbf{Q}[\mathbf{m}_h^n]$ from (18) in $\tilde{\mathbf{P}}_Q[\mathbf{m}_h^n]$ has a block-diagonal form and is explicitly available at each time-step. Hence, \mathbf{P}^{2D} and $\tilde{\mathbf{P}}_Q[\mathbf{m}_h^n]$ have a similar computational complexity. In contrast to that, the theoretical preconditioner $\mathbf{P}_Q[\mathbf{m}_h^n]$ from (23) has to be rebuilt at every time-step.

5.1. An academic example. We investigate the iteration number for different choices \mathbf{P} in (23) with respect to the mesh-size h . To this end, we adapt the setting of [PRS18, Section 6.1]: On $\Omega := (0, 1)^3$, we employ the initial value and applied field

$$\mathbf{m}^0 := (1, 0, 0)^T \quad \text{and} \quad \mathbf{f}(x_1, x_2, x_3) := 10 (\sin(x_1), \cos(x_1), 0)^T,$$

respectively. We set $T = 1$, $\alpha = 0.5$, $\ell_{\text{ex}}^2 = 10 \equiv \beta(k)$; see also (10b). The lower-order contributions $\boldsymbol{\pi}$ consist only of the stray field. With this configuration, the magnetization is expected to align itself in the direction of (the non-constant) applied field \mathbf{f} . For time discretization, we fix $k = 10^{-2}$. For space discretization, we employ the meshes generated by the NGS/Py [ngs] embedded module `Netgen` with the mesh-sizes

$$h \in \left\{ \frac{j}{2} \cdot 10^{-2} : j = 5, 6, \dots, 20 \right\}.$$

We employ the different preconditioners from Section 4 with $\alpha_{\mathbf{P}} = 1$ or do not use preconditioning for the iterative solution of the underlying linear system (19). Moreover, we always fix $\mathbf{T}_n := \mathbf{I}_{3 \times 3}$, i.e., we always employ the standard choice (17b).

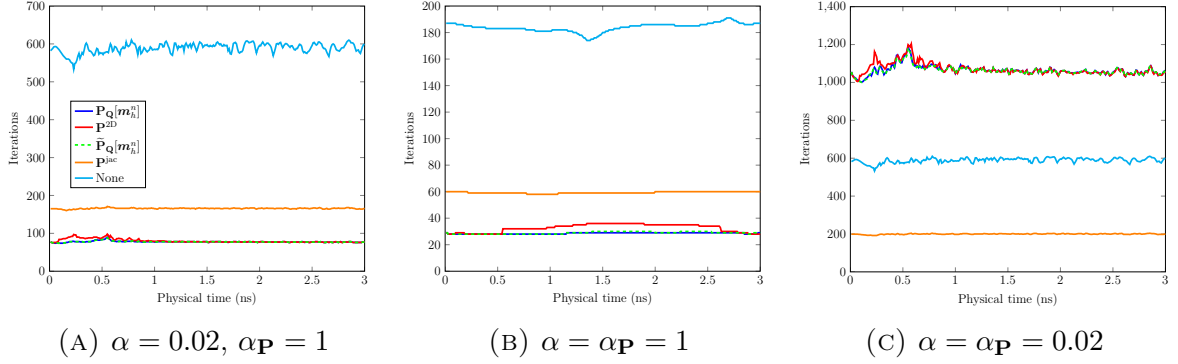


FIGURE 2. Experiment of Section 5.2: GMRES iterations over time.

In Figure 1, we plot the average number of GMRES iterations. As expected, no preconditioning (None) requires the most iterations. The Jacobi preconditioner \mathbf{P}^{jac} brings a slight improvement. However, (None) and \mathbf{P}^{jac} are both not robust with respect to the mesh-size h . In contrast to that, the theoretical ($\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$), the stationary ($\mathbf{P}^{2\text{D}}$), and the practical ($\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$) preconditioner from Section 4.1–4.3, respectively, require significantly less iterations. Moreover, all three latter options are robust with respect to the mesh-size h . However, $\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$ is computationally more expensive than $\mathbf{P}^{2\text{D}}$ and $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$; cf. Remark 9. In conclusion, the latter experiment suggests to use either the stationary preconditioner $\mathbf{P}^{2\text{D}}$ or the practical preconditioner $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$.

5.2. μMAG standard problem #4. We investigate the practical applicability of the different choices for the preconditioner \mathbf{P} from Section 4 to a physically relevant example. To this end, we consider the $\mu\text{-MAG}$ standard problem #4 [mum], which simulates the switching of the magnetization in a thin permalloy layer

$$\Omega := (-250\text{nm}, 250\text{nm}) \times (-62.5\text{nm}, 62.5\text{nm}) \times (-1.5\text{nm}, 1.5\text{nm}).$$

The dynamics is described by LLG (stated in physical SI units), which reads

$$\partial_t \mathbf{M} = -\gamma_0 \mathbf{M} \times \mathbf{H}_{\text{eff}}(\mathbf{M}) + \frac{\alpha}{M_s} \mathbf{M} \times \partial_t \mathbf{M} \quad \text{in } (0, T) \times \Omega =: \Omega_T, \quad (43a)$$

$$\partial_{\mathbf{n}} \mathbf{M} = \mathbf{0} \quad \text{on } (0, T) \times \partial\Omega, \quad (43b)$$

$$\mathbf{M}(0) = \mathbf{M}^0 \quad \text{in } \Omega. \quad (43c)$$

Here, $\gamma_0 = 2.21 \cdot 10^5 \text{N/A}^2$ is the gyromagnetic ratio $M_s := 8.0 \cdot 10^5 \text{A/m}$ is the saturation magnetization, and the sought magnetization \mathbf{M} satisfies $|\mathbf{M}| = M_s$ a.e. in Ω_T . Moreover, we employ the physical end time $T := 3\text{ns}$. The physical effective field reads

$$\mathbf{H}_{\text{eff}}(\mathbf{M}) := \frac{2A}{\mu_0 M_s^2} \Delta \mathbf{M} + \boldsymbol{\pi}(\mathbf{M}) + \mathbf{H}_{\text{ext}}. \quad (43d)$$

Here, $A := 1.3 \cdot 10^{-11} \text{J/m}$ is the exchange constant of permalloy, $\mu_0 = 4\pi \cdot 10^{-7} \text{N/A}^2$ is the permeability of vacuum, $\boldsymbol{\pi}(\mathbf{M})$ consists only of the stray field, and

$$\mu_0 \mathbf{H}_{\text{ext}} := (-35.5, -6.3, 0)^T \text{mT} \quad (43e)$$

is a constant applied field corresponding to the second $\mu\text{-MAG}$ #4 configuration.

Our numerical simulation is based on the non-dimensional form (1) of (43): With $L = 10^{-9}$ and the temporal and spatial rescaling $t \mapsto (\gamma_0 M_s t)/L$ and $\mathbf{x} \mapsto \mathbf{x}/L$, respectively, the dimensionless unknown $\mathbf{m} := \mathbf{M}/M_s$ satisfies LLG (1) in the (not relabelled) domain

$$\Omega := (-250, 250) \times (-62.5, 62.5) \times (-1.5, 1.5),$$

and the parameters

$$\ell_{\text{ex}}^2 = \frac{2A}{\mu_0 M_s^2 L^2} \approx 32.3283 \quad \text{and} \quad \mathbf{f} := (-28250, -5013.4, 0)^T. \quad (44)$$

For time and space discretization, we choose the physical time-step size $\Delta t = 0.1\text{ps}$ and the physical mesh-size $\Delta \mathbf{x} = 5\text{nm}$. With the above rescaling, we use Algorithm 1 with the actual numerical discretization parameters

$$k = \gamma_0 M_s \Delta t \approx 0.017688, \quad \text{and} \quad h = \Delta x/L = 5. \quad (45)$$

We employ the corresponding mesh generated by the NGS/Py [ngs] embedded module `Netgen`, which consists of 17478 elements and 6073 nodes. We employ the preconditioners from Section 4 or do not use preconditioning for the iterative solution of the underlying linear system (19). Moreover, we employ the adaptive strategy for \mathbf{T}_n from Section 4.5.

5.2.1. μ -MAG #4 configuration. As specified by μ -MAG #4, we choose $\alpha = 0.02$ for the Gilbert damping constant. For preconditioning, we choose $\alpha_{\mathbf{P}} = 1$. In Figure 2, we plot the required GMRES iterations for different preconditioners over time. No preconditioning (None) requires the most iterations. The theoretical ($\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$), the stationary ($\mathbf{P}^{2\text{D}}$), the practical ($\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$), and the Jacobi preconditioner (\mathbf{P}^{jac}) significantly improve the iteration number. Here, $\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$, $\mathbf{P}^{2\text{D}}$, and $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ require the fewest iterations and $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ performs even slightly better than $\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$ and $\mathbf{P}^{2\text{D}}$. Moreover, recall that $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ and $\mathbf{P}^{2\text{D}}$ have similar computational complexity, while $\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$ has to be rebuilt at every time-step; cf. Remark 9. This experiment suggests either the stationary preconditioner $\mathbf{P}^{2\text{D}}$ or the practical preconditioner $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$. In the following three subsections, we extend this experiment to discuss the roles of α (see Section 5.2.2), \mathbf{T}_n (see Section 5.2.4), and $\alpha_{\mathbf{P}}$ (see Section 5.2.3).

5.2.2. The impact of α . We repeat the experiment from Section 5.2.1 with $\alpha = \alpha_{\mathbf{P}} = 1$. The new setting still simulates the switching dynamics of μ -MAG #4, but with a bigger Gilbert damping constant α . In Figure 2, we plot the required GMRES iterations for different preconditioners over time. Compared to the original Figure 5.2.1 with $\alpha = 0.02$ and $\alpha_{\mathbf{P}} = 1$, all approaches require less iterations. This is in accordance with the results from Section 4, where larger α leads to better contraction of the residuals; cf., e.g., (25).

5.2.3. The impact of $\alpha_{\mathbf{P}}$. First, we repeat the experiment with $\alpha_{\mathbf{P}} = \alpha = 0.02$. Unlike Section 5.2.1 and Section 5.2.2, we observe the following: In Figure 2, we plot the required GMRES iterations for different preconditioners over time. The Jacobi-preconditioner (\mathbf{P}^{jac}) requires less iterations than no preconditioning (None). The theoretical ($\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$), the stationary ($\mathbf{P}^{2\text{D}}$), and the practical preconditioner ($\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$) fail completely. All three require significantly more GMRES iterations than no preconditioning (None). This might be due to the skew-symmetric part $\mathbf{S}[\mathbf{m}_h^n]$ in the (yet unconstrained system) matrix $\mathbf{A}_k[\mathbf{m}_h^n]$ from (14). Here, $\mathbf{S}[\mathbf{m}_h^n]$ is similar to a mass matrix, but unlike $\mathbf{M}_k[\mathbf{m}_h^n]$, it lacks

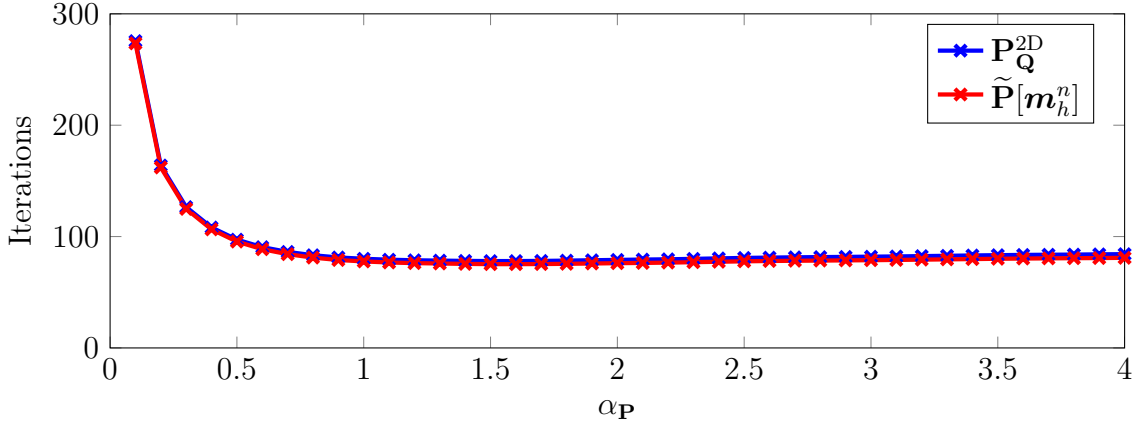


FIGURE 3. Experiment of Section 5.2.3: Average number of GMRES iterations ($\alpha = 0.02$).

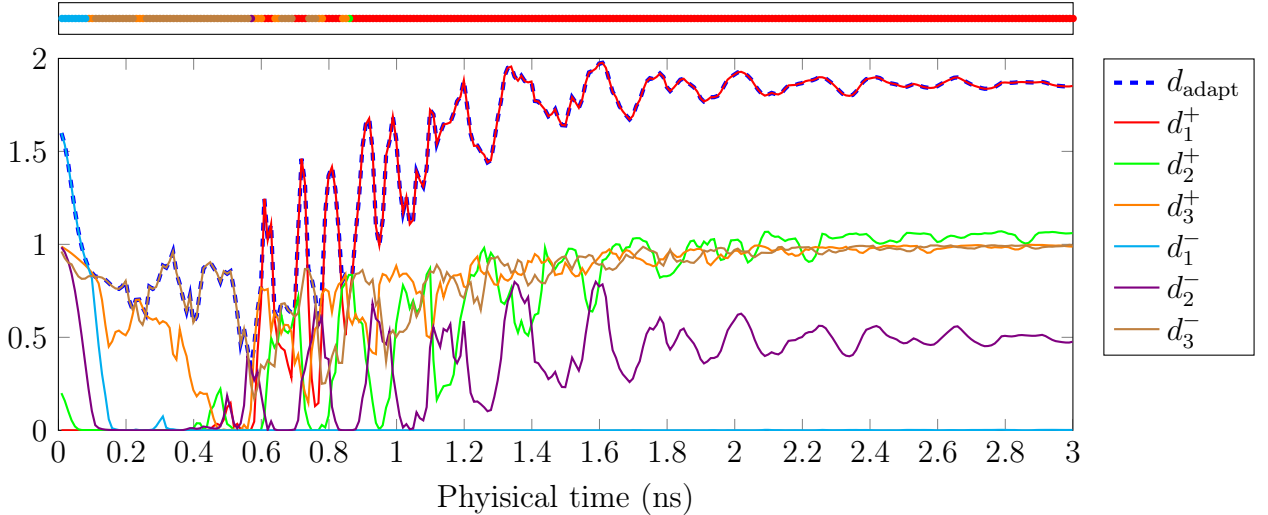


FIGURE 4. Experiment of Section 5.2.4: Evolution of d_ℓ^* with $\ell \in \{1, 2, 3\}$ and $\star \in \{+, -\}$, and $d_{\text{adapt}} := \max_{\ell \in \{1, 2, 3\}} \{d_\ell^+, d_\ell^-\}$ over time ($\alpha = 0.02$, $\alpha_{\mathbf{P}} = 1$). The above line indicates the current state of d_{adapt} .

the factor α . To empirically determine a good choice for $\alpha_{\mathbf{P}}$, we repeat our current experiment with the fixed preconditioners $\mathbf{P}^{2\text{D}}$ and $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$, and vary $\alpha_{\mathbf{P}}$. In Figure 3, we plot the average number of the required GMRES iterations. As already observed in Figure 2, bigger values of $\alpha_{\mathbf{P}}$ result in significantly less iterations. However, $\alpha_{\mathbf{P}}$ bigger than 1, has little to no effect. In conclusion, we suggest to always choose $\alpha_{\mathbf{P}} = 1$.

5.2.4. *Adaptive vs. fixed \mathbf{T}_n .* We extend the experiment from Section 5.2.1 to discuss the impact of the adaptive strategy for $\mathbf{T}_n \in \mathbb{R}^{3 \times 3}$ from Section 4.5. In Figure 4, we plot the evolution of $d_\ell^* \in [0, 2]$ from (40), where $\ell \in \{1, 2, 3\}$ and $\star \in \{+, -\}$ and $d_{\text{adapt}} := \max_{\ell \in \{1, 2, 3\}} \{d_\ell^+, d_\ell^-\}$. Recall from (42) that we can choose $\gamma = d_{\text{adapt}}$ for adaptive

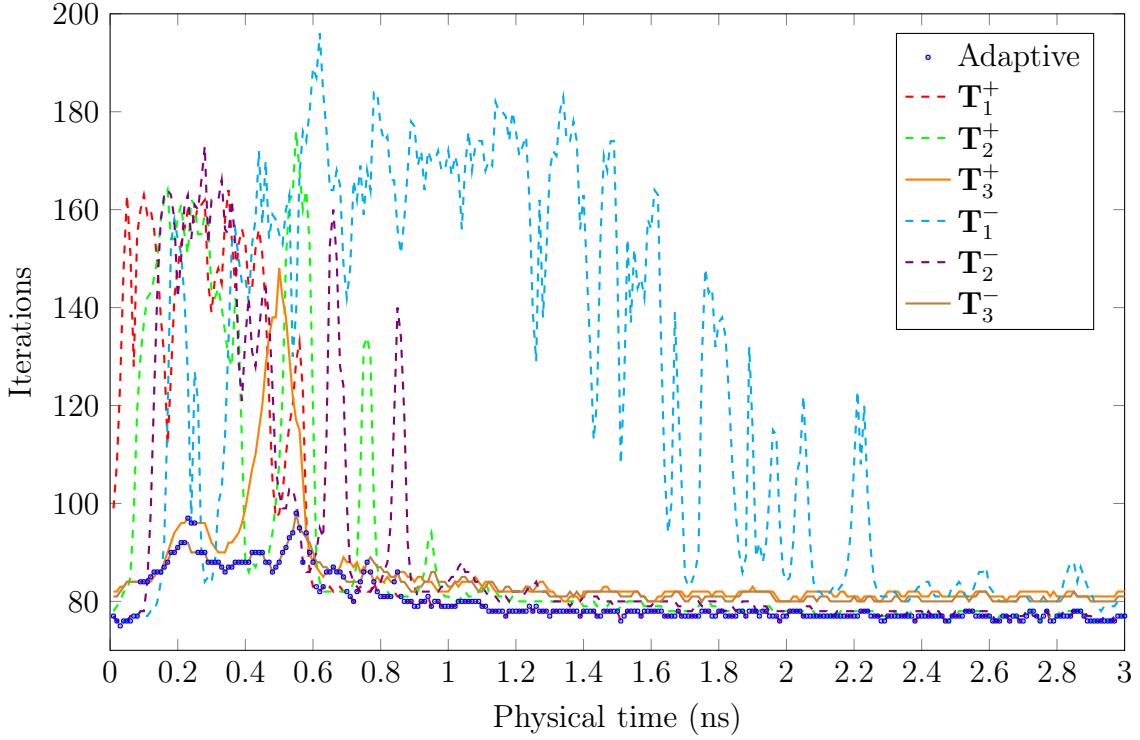


FIGURE 5. Experiment of Section 5.2.4: GMRES iterations for the stationary preconditioner \mathbf{P}^{2D} over time ($\alpha = 0.02$, $\alpha_{\mathbf{P}} = 1$).

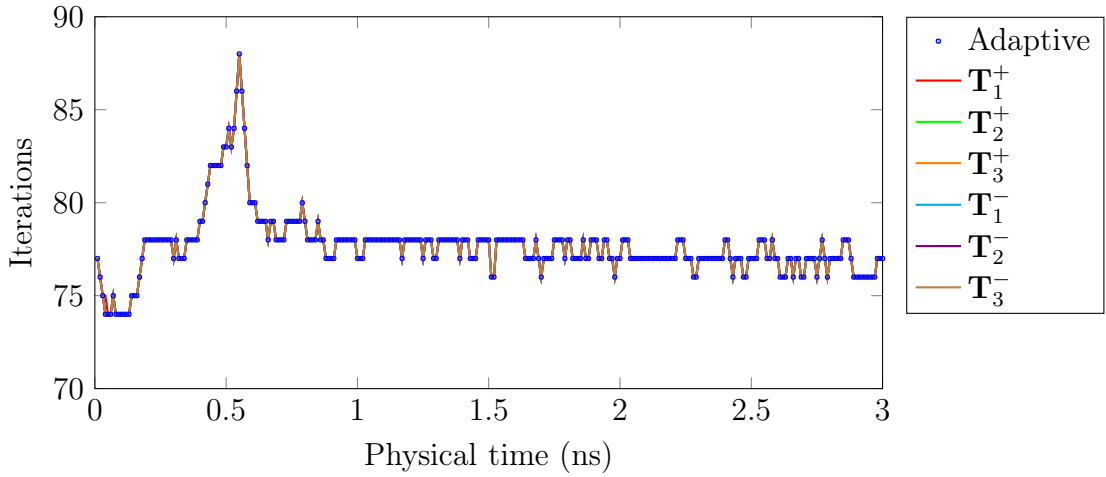


FIGURE 6. Experiment of Section 5.2.4: GMRES iterations for the practical preconditioner $\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ over time ($\alpha = 0.02$, $\alpha_{\mathbf{P}} = 1$).

\mathbf{T}_n and $\gamma = d_\ell^*$ for fixed $\mathbf{T}_n := \mathbf{T}_\ell^*$. For adaptive \mathbf{T}_n , we always have in our example that $1 + (\mathbf{T}_n \mathbf{m}_h^n(z))_3 \geq \gamma > 0$, i.e., Corollary 6 and Theorem 7 (ii) apply.

In Figure 5, we consider the stationary preconditioner \mathbf{P}^{2D} . We plot the evolution of the GMRES iteration numbers with the corresponding \mathbf{T}_n (Adaptive) as well as fixed

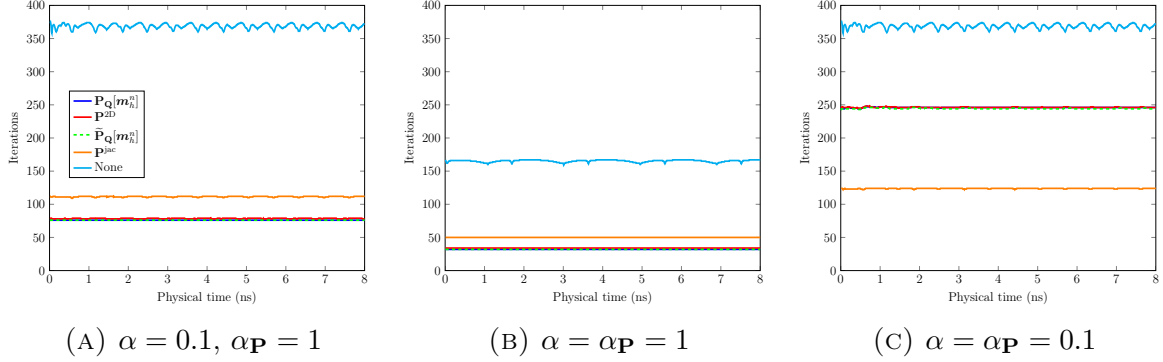


FIGURE 7. Experiment of Section 5.3: GMRES iterations over time.

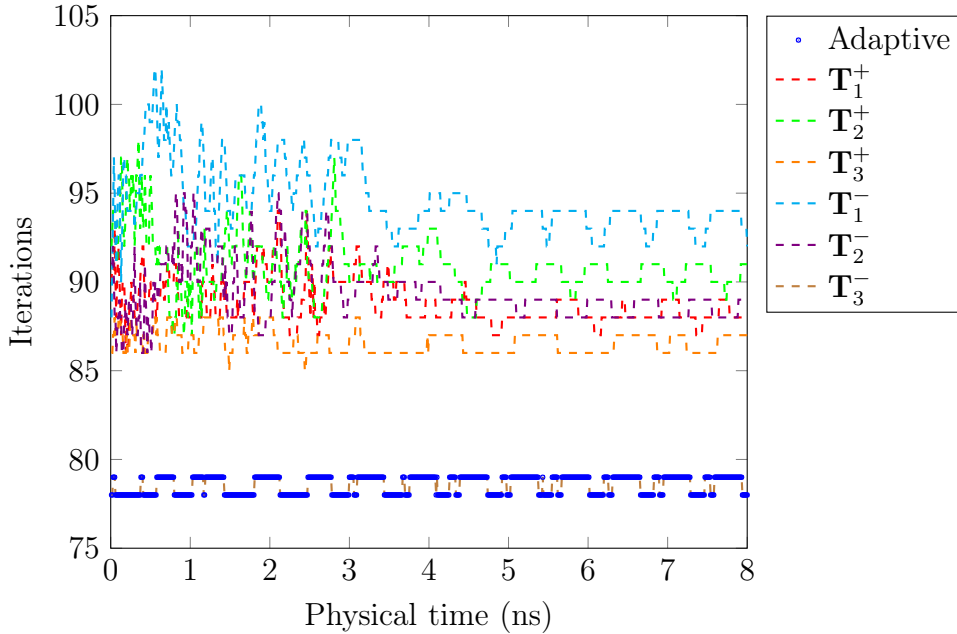


FIGURE 8. Experiment of Section 5.3: GMRES iterations for the stationary preconditioner \mathbf{P}^{2D} over time ($\alpha = 0.1, \alpha_{\mathbf{P}} = 1$).

$\mathbf{T}_n := \mathbf{T}_\ell^\star$ from (41), where $\ell \in \{1, 2, 3\}$ and $\star \in \{+, -\}$. In Figure 6, we repeat this experiment with the practical preconditioner $\tilde{\mathbf{P}}_{\mathbf{Q}}[m_h^n]$ instead of \mathbf{P}^{2D} .

Adaptive \mathbf{T}_n is not always the best choice, however, it avoids the increased iteration number of a fixed \mathbf{T}_n ; see Figure 5. Yet, for the relation of the iteration number of fixed \mathbf{T}_n and the corresponding d_ℓ^\star , the picture is not complete: In Figure 6, all options appear to be equally good, even though, e.g., for fixed $\mathbf{T}_n := \mathbf{T}_1^-$, it holds that $d_1^- \approx 0$ most of the time; see Figure 4. In conclusion, our experiment suggests to use adaptive \mathbf{T}_n . However, a full understanding of the effect of the choice of \mathbf{T}_n might require further work.

5.3. μ MAG standard problem #5 . As for μ -MAG #4 in Section 5.2, the the μ -MAG standard problem #5 [mum] relies on a physical formulation similar to (43) with the end time $T = 8\text{ns}$ and the domain $\Omega := (-50\text{nm}, 50\text{nm}) \times (-50\text{nm}, 50\text{nm}) \times (-5\text{nm}, 5\text{nm})$.

The simulation is based on the non-dimensional form (1), where we employ the same temporal and spatial rescaling as in Section 5.2 with the saturation magnetization $M_s := 8.0 \cdot 10^5 \text{ A/m}$. In particular, we compute ℓ_{ex}^2 as in (44), where $A := 1.3 \cdot 10^{-11} \text{ J/m}$, $\mu_0 = 4\pi \cdot 10^{-7} \text{ N/A}^2$, and $\alpha := 0.1$. In extension to $\mu\text{Mag} \#4$, the $\mu\text{-MAG}$ standard problem $\#5$ involves LLG (1) with an extended effective field

$$\mathbf{h}_{\text{eff}}(\mathbf{m}) := \ell_{\text{ex}}^2 \Delta \mathbf{m} + \boldsymbol{\pi}(\mathbf{m}) + \mathbf{f} + \mathbf{m} \times (\mathbf{u} \cdot \nabla) \mathbf{m} + \beta (\mathbf{u} \cdot \nabla) \mathbf{m},$$

where the additional term is the Zhang-Li spin-torque term [ZL04, TNMS05]. Here, $\mathbf{f} := \mathbf{0}$, $\boldsymbol{\pi}(\mathbf{m})$ is the stray field, $\beta := 0.05$ is the constant of non-adiabacity, and $\mathbf{u} := 1/(\gamma_0 M_s L)(72.17, 0, 0)^T$ is the spin current with the gyromagnetic ratio $\gamma_0 = 2.21 \cdot 10^5 \text{ N/A}^2$ and the scaling parameter $L = 10^{-9}$.

For time and space discretization, we choose the physical time-step size $\Delta t = 0.1 \text{ ps}$ and the physical mesh-size $\Delta x = 3 \text{ nm}$. The corresponding k and h are obtained as in (45). The corresponding mesh generated by the NGS/Py [ngs] embedded module Netgen has 25666 elements and 5915 nodes. We employ our C++-code, the first order tangent plane scheme (see (10)). GMRES is restarted every 200 iterations and an iteration tolerance $\varepsilon = 10^{-14}$. We perform the same numerical experiments as for $\mu\text{MAG} \#4$. The results qualitatively agree with those of Section 5.2; see in Figure 7 and Figure 8.

6. PROOF OF MAIN RESULTS

6.1. Auxiliary mappings. Recall the hat functions $(\varphi_i)_{i=1}^N$, the nodal basis $(\boldsymbol{\psi}_i)_{i=1}^{2N}$ of $(\mathcal{S}_h)^2$ from (16), and $(\boldsymbol{\phi}_i)_{i=1}^{3N}$ of $\mathcal{S}_h = (\mathcal{S}_h)^3$ from (12). Recall the definitions of $\mathbf{H}[\cdot] \in \mathbb{R}^{3 \times 2}$ from (17) and $\mathbf{Q}[\cdot] \in \mathbb{R}^{3N \times 2N}$ from (18). Given $\boldsymbol{\mu}_h \in \mathcal{M}_h$, define the mappings

$$\mathbb{P}_h[\boldsymbol{\mu}_h] : (\mathcal{S}_h)^2 \rightarrow \mathcal{K}_h[\boldsymbol{\mu}_h] \subsetneq \mathcal{S}_h : \mathbf{w}_h \mapsto \sum_{i=1}^N (\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_i)] \mathbf{w}_h(z_i)) \varphi_i, \quad (46a)$$

$$\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] : \mathbb{R}^{2N} \rightarrow \mathcal{K}_h[\boldsymbol{\mu}_h] \subsetneq \mathcal{S}_h : \mathbf{x} \mapsto \sum_{i=1}^{3N} (\mathbf{Q}[\boldsymbol{\mu}_h] \mathbf{x})_i \boldsymbol{\phi}_i, \quad (46b)$$

their ‘‘transposed’’ versions

$$\mathbb{P}_h^T[\boldsymbol{\mu}_h] : \mathcal{S}_h \rightarrow (\mathcal{S}_h)^2 : \mathbf{v}_h \mapsto \sum_{i=1}^N (\mathbf{H}[\boldsymbol{\mu}_h(z_i)]^T \mathbf{T}_n \mathbf{v}_h(z_i)) \varphi_i \quad (47a)$$

$$\tilde{\mathbb{P}}_h^T[\boldsymbol{\mu}_h] : \mathbb{R}^{3N} \rightarrow (\mathcal{S}_h)^2 : \mathbf{x} \mapsto \sum_{i=1}^{3N} (\mathbf{Q}[\boldsymbol{\mu}_h]^T \mathbf{x})_i \boldsymbol{\phi}_i, \quad (47b)$$

and the compositions

$$\boldsymbol{\Pi}_h[\boldsymbol{\mu}_h] : \mathcal{S}_h \rightarrow \mathcal{K}_h[\boldsymbol{\mu}_h] \subsetneq \mathcal{S}_h : \mathbf{v}_h \mapsto \sum_{i=1}^N (\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_i)] \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_i)]^T \mathbf{T}_n \mathbf{v}_h(z_i)) \varphi_i, \quad (48a)$$

$$\tilde{\boldsymbol{\Pi}}_h[\boldsymbol{\mu}_h] : \mathbb{R}^{3N} \rightarrow \mathcal{K}_h[\boldsymbol{\mu}_h] \subsetneq \mathcal{S}_h : \mathbf{x} \mapsto \sum_{i=1}^{3N} (\mathbf{Q}[\boldsymbol{\mu}_h] \mathbf{Q}[\boldsymbol{\mu}_h]^T \mathbf{x})_i \boldsymbol{\phi}_i. \quad (48b)$$

Note that $\mathbf{\Pi}_h[\boldsymbol{\mu}_h]$ is the nodewise orthogonal projection onto $\mathcal{K}_h[\boldsymbol{\mu}_h]$. The following lemma discusses the relations of the mappings (46)–(48).

Lemma 10. *For any $\boldsymbol{\mu}_h \in \mathcal{M}_h$, there hold the following assertions (i)–(v):*

- (i) *For $\mathbf{x} \in \mathbb{R}^{2N}$ and $\mathbf{w}_h := \sum_{i=1}^{2N} \mathbf{x}_i \boldsymbol{\psi}_i \in (\mathcal{S}_h)^2$, it holds that $\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h = \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}$.*
 - (ii) *For $\mathbf{y} \in \mathbb{R}^{3N}$ and $\mathbf{v}_h := \sum_{i=1}^{3N} \mathbf{y}_i \boldsymbol{\phi}_i$, it holds that $\mathbf{\Pi}_h[\boldsymbol{\mu}_h] \mathbf{v}_h = \tilde{\mathbf{\Pi}}_h[\boldsymbol{\mu}_h] \mathbf{y}$.*
 - (iii) *For $\mathbf{v}_h \in \mathcal{S}_h$, it holds that $\mathbb{P}_h[\boldsymbol{\mu}_h] \circ \mathbb{P}_h^T[\boldsymbol{\mu}_h] \mathbf{v}_h = \mathbf{\Pi}_h[\boldsymbol{\mu}_h] \mathbf{v}_h$.*
 - (iv) *For $\mathbf{w}_h \in (\mathcal{S}_h)^2$, it holds that $\mathbb{P}_h^T[\boldsymbol{\mu}_h] \circ \mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h = \mathbf{w}_h$.*
 - (v) *$\mathbb{P}_h[\boldsymbol{\mu}_h] : (\mathcal{S}_h)^2 \rightarrow \mathcal{K}_h[\boldsymbol{\mu}_h]$ and $\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] : \mathbb{R}^{2N} \rightarrow \mathcal{K}_h[\boldsymbol{\mu}_h]$ are isomorphisms.*
-

Proof. (i)–(ii) follow by definition. (iii)–(iv) follow from the block-structure of $\mathbf{Q}[\boldsymbol{\mu}_h]$, since $\mathbf{Q}[\boldsymbol{\mu}_h]$ has orthonormal columns. Since $(\mathcal{S}_h)^2 \cong \mathbb{R}^{2N} \cong \mathcal{K}_h[\boldsymbol{\mu}_h]$, (iv) proves that $\mathbb{P}_h[\boldsymbol{\mu}_h]$ is an isomorphism. Together with (i), this also proves the statement about $\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h]$. Altogether, this concludes the proof. \square

In the following lemma, we prove discrete \mathbf{L}^2 -stabilities of the mappings (46)–(48).

Lemma 11. *There exists $C > 0$, which depends only on C_{mesh} , such that the following assertions (i)–(vi) hold true:*

- (i) *For any $\boldsymbol{\mu}_h \in \mathcal{M}_h$, it holds that*

$$C^{-1} \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} \leq \|\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} \leq C \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} \quad \text{for all } \mathbf{w}_h \in (\mathcal{S}_h)^2.$$

- (ii) *For any $\boldsymbol{\mu}_h \in \mathcal{M}_h$, it holds that*

$$\|\mathbb{P}_h^T[\boldsymbol{\mu}_h] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)} \leq C \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)} \quad \text{for all } \mathbf{v}_h \in \mathcal{S}_h.$$

- (iii) *For any $\boldsymbol{\mu}_h \in \mathcal{M}_h$, it holds that*

$$\|\mathbf{\Pi}_h[\boldsymbol{\mu}_h] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)} \leq C \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)} \quad \text{for all } \mathbf{v}_h \in \mathcal{S}_h.$$

- (iv) *For any $\boldsymbol{\mu}_h \in \mathcal{M}_h$, it holds that*

$$C^{-1} h^{3/2} |\mathbf{x}| \leq \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}\|_{\mathbf{L}^2(\Omega)} \leq C h^{3/2} |\mathbf{x}| \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

- (v) *For any $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h \in \mathcal{M}_h$, it holds that*

$$C^{-1} \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{\mathbf{L}^2(\Omega)} \leq \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}\|_{\mathbf{L}^2(\Omega)} \leq C \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{\mathbf{L}^2(\Omega)} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

- (vi) *For any $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h \in \mathcal{M}_h$, it holds that*

$$\|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} - \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{\mathbf{L}^2(\Omega)} \leq C h^{3/2} \max_{z \in \mathcal{N}_h} \|\mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z)] - \mathbf{H}[\mathbf{T}_n \boldsymbol{\nu}_h(z)]\| |\mathbf{x}| \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

Proof. Throughout the proof, recall that $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$. For the proof of (i), and (iv)–(vi), let $\mathbf{x} \in \mathbb{R}^{2N}$ and define $\mathbf{w}_h := \sum_{i=1}^{2N} \mathbf{x}_i \boldsymbol{\psi}_i$.

The matrices $\mathbf{H}[\boldsymbol{\mu}_h(z_i)] \in \mathbb{R}^{3 \times 2}$ have orthonormal columns. Lemma 10 (i) yields that

$$\begin{aligned} \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}\|_{\mathbf{L}^2(\Omega)}^2 &= \|\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h\|_{\mathbf{L}^2(\Omega)}^2 \\ &\stackrel{(5)}{\simeq} h^3 \sum_{i=1}^N |\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_i)] \mathbf{w}_h(z_i)|^2 = h^3 \sum_{i=1}^N |\mathbf{w}_h(z_i)|^2 \stackrel{(5)}{\simeq} \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)}^2. \end{aligned}$$

This proves (i), and (iv) follows from $\sum_{i=1}^N |\mathbf{w}(z_i)|^2 = \sum_{i=1}^{2N} |\mathbf{x}_i|^2$. (v) is a direct consequence of (iv). For the proof of (vi), note that

$$\begin{aligned} \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h]\mathbf{x} - \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h]\mathbf{x}\|_{\mathbf{L}^2(\Omega)}^2 &\stackrel{(5)}{\simeq} h^3 \sum_{i=1}^N \left| \mathbf{T}_n(\mathbf{H}[\mathbf{T}_n\boldsymbol{\mu}_h(z_i)] - \mathbf{H}[\mathbf{T}_n\boldsymbol{\nu}_h(z_i)]) \mathbf{w}_h(z_i) \right|^2 \\ &\leq h^3 \max_{z \in \mathcal{N}_h} \|\mathbf{H}[\mathbf{T}_n\boldsymbol{\mu}_h(z)] - \mathbf{H}[\mathbf{T}_n\boldsymbol{\nu}_h(z)]\|^2 |\mathbf{x}|^2. \end{aligned}$$

This proves (vi). For the proof of (ii)–(iii), let $\mathbf{v}_h \in \mathcal{S}_h$. Since the matrices $\mathbf{H}[\boldsymbol{\mu}_h(z_i)] \in \mathbb{R}^{3 \times 2}$ have orthonormal columns, we obtain that

$$\|\mathbb{P}_h^T[\boldsymbol{\mu}_h]\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 \stackrel{(5)}{\simeq} h^3 \sum_{i=1}^N |\mathbf{H}[\mathbf{T}_n\boldsymbol{\mu}_h(z_i)]^T \mathbf{T}_n \mathbf{v}_h(z_i)|^2 \leq h^3 \sum_{i=1}^N |\mathbf{v}_h(z_i)|^2 \stackrel{(5)}{\simeq} \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2.$$

This proves (ii). Together with (i)–(ii) and Lemma 10 (iii), this also proves (iii). Altogether, this concludes the proof. \square

In the following lemma, we prove discrete \mathbf{H}^1 -stability properties of the mappings (46)–(48). Note that (in contrast to Lemma 10 and Lemma 11) the following lemma builds on the explicit definition of the Householder matrices (17).

Lemma 12. *Let $\boldsymbol{\mu}_h \in \mathcal{M}_h$ with $1 + (\boldsymbol{\mu}_h(z))_3 \geq \gamma > 0$ for all $z \in \mathcal{N}_h$. Then, there exists $C > 1$, which depends only on C_{mesh} , such that the following assertions (i)–(iii) hold true:*

(i) *For all $\mathbf{w}_h \in (\mathcal{S}_h)^2$, it holds that*

$$\|\nabla(\mathbb{P}_h[\boldsymbol{\mu}_h]\mathbf{w}_h)\|_{\mathbf{L}^2(\Omega)} \leq C\gamma^{-2} \|\nabla\boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} + C \|\nabla\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)}.$$

(ii) *For all $\mathbf{v}_h \in \mathcal{S}_h$, it holds that*

$$\|\nabla(\mathbb{P}_h^T[\boldsymbol{\mu}_h]\mathbf{v}_h)\|_{\mathbf{L}^2(\Omega)} \leq C\gamma^{-2} \|\nabla\boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)} + C \|\nabla\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}.$$

(iii) *For all $\mathbf{v}_h \in \mathcal{S}_h$, it holds that*

$$\|\nabla(\mathbf{\Pi}_h[\boldsymbol{\mu}_h]\mathbf{v}_h)\|_{\mathbf{L}^2(\Omega)} \leq C\gamma^{-2} \|\nabla\boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)} + C \|\nabla\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}.$$

Proof. First, we prove (i). We split the proof into the following six steps.

Step 1. We derive a handier representation of $\mathbb{P}_h[\boldsymbol{\mu}_h]$. Let $\mu_1, \mu_2, \mu_3 \in \mathcal{S}_h$ such that $\mathbf{T}_n\boldsymbol{\mu}_h := (\mu_1, \mu_2, \mu_3)^T$. Since functions in \mathcal{S}_h attain their minimum in one of the nodes, we obtain, in particular, that

$$1 + \mu_3 = 1 + (\mathbf{T}_n\boldsymbol{\mu}_h)_3 \geq \gamma > 0 \quad \text{in } \Omega. \quad (49)$$

Hence, we can interpret

$$\mathbf{R}_1[\mathbf{T}_n\boldsymbol{\mu}_h] := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -\mu_1 & -\mu_2 \end{pmatrix}, \quad \mathbf{R}_2[\mathbf{T}_n\boldsymbol{\mu}_h] := \frac{1}{1 + \mu_3} \begin{pmatrix} \mu_1^2 & \mu_1\mu_2 \\ \mu_1\mu_2 & \mu_2^2 \\ 0 & 0 \end{pmatrix} \quad (50)$$

and

$$\mathbf{q}[\boldsymbol{\mu}_h] := \mathbf{T}_n\mathbf{R}_1[\mathbf{T}_n\boldsymbol{\mu}_h] - \mathbf{T}_n\mathbf{R}_2[\mathbf{T}_n\boldsymbol{\mu}_h] \quad (51)$$

as functions $\mathbf{q}[\boldsymbol{\mu}_h], \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h], \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] : \Omega \rightarrow \mathbb{R}^{3 \times 2}$. With the definition of the Householder matrices (17), an elementary calculation shows that $\mathbf{T}_n \mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z_i)] = \mathbf{q}[\boldsymbol{\mu}_h(z_i)]$ for all $i = 1, \dots, N$. With \mathcal{I}_h being the vector-valued nodal interpolant onto \mathcal{S}_h , we get

$$\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h \stackrel{(46)}{=} \mathcal{I}_h(\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h) \quad \text{for all } \mathbf{w}_h \in (\mathcal{S}_h)^2. \quad (52)$$

Step 2. We derive preliminary estimates for $\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]$ and $\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]$ from (50). To this end, recall that $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$. Lemma A.1 (i) yields that

$$\|\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]\|_{\mathbf{L}^\infty(\Omega)} \lesssim 1 \quad \text{and} \quad \|\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]\|_{\mathbf{L}^\infty(\Omega)} \lesssim 1. \quad (53a)$$

Let $k \in \{1, 2, 3\}$. Lemma A.1 (ii) yields that

$$\|\partial_k [\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]]\|_{\mathbf{L}^\infty(\Omega)} \stackrel{(49)}{\lesssim} \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} = \|\mathbf{T}_n \nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} = \|\nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}, \quad (53b)$$

as well as

$$\|\partial_k [\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]]\|_{\mathbf{L}^\infty(\Omega)} \stackrel{(49)}{\lesssim} \gamma^{-1} \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} = \gamma^{-1} \|\nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}. \quad (53c)$$

Let $\ell, k \in \{1, 2, 3\}$. The definition (50) and Lemma A.1 (iii) yield that, elementwise,

$$\partial_\ell \partial_k [\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]] = \mathbf{0} \quad (53d)$$

as well as

$$\max_{K \in \mathcal{T}_h} \|\partial_\ell \partial_k [\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]]\|_{\mathbf{L}^\infty(K)} \stackrel{(49)}{\lesssim} \gamma^{-2} \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}^2 = \gamma^{-2} \|\nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)}^2. \quad (53e)$$

Step 3. For $\mathbf{w}_h \in (\mathcal{S}_h)^2$, we estimate $\|\nabla [\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h]\|_{\mathbf{L}^2(\Omega)}$. To that end, note that $\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h|_K \in (H^2(K))^3$ for all elements K . We exploit the elementwise approximation properties of the nodal interpolant \mathcal{I}_h and obtain that

$$\begin{aligned} \|\nabla [\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h]\|_{\mathbf{L}^2(\Omega)} &\stackrel{(46)}{=} \|\nabla [\mathcal{I}_h(\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h)]\|_{\mathbf{L}^2(\Omega)} \\ &\lesssim \|\nabla [\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h]\|_{\mathbf{L}^2(\Omega)} + \|\nabla [(1 - \mathcal{I}_h)(\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h)]\|_{\mathbf{L}^2(\Omega)} \\ &\lesssim \|\nabla [\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h]\|_{\mathbf{L}^2(\Omega)} + h \left(\sum_{K \in \mathcal{T}_h} \|D^2[(\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h)]\|_{\mathbf{L}^2(K)}^2 \right)^{1/2} =: T_1 + h T_2. \end{aligned} \quad (54)$$

Step 4. We estimate T_1 . Let $k \in \{1, 2, 3\}$. The product rule yields that

$$\begin{aligned} \partial_k [\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h] &= \partial_k [\mathbf{q}[\boldsymbol{\mu}_h]] \mathbf{w}_h + \mathbf{q}[\boldsymbol{\mu}_h] \partial_k \mathbf{w}_h \\ &\stackrel{(51)}{=} \partial_k [\mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]] \mathbf{w}_h - \partial_k [\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \mathbf{w}_h + \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] \partial_k \mathbf{w}_h - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] \partial_k \mathbf{w}_h \\ &= \mathbf{T}_n \partial_k [\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]] \mathbf{w}_h - \mathbf{T}_n \partial_k [\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \mathbf{w}_h + \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] \partial_k \mathbf{w}_h - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] \partial_k \mathbf{w}_h. \end{aligned} \quad (55)$$

Note that $\gamma \leq 1 + \mu_3 \leq 1 + |\mu_3| \leq 2$ and recall that $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$. With the estimates from (53) and with $1 \leq 2/\gamma$, the latter equation yields that

$$\begin{aligned} T_1 &\stackrel{(54)}{\lesssim} \gamma^{-1} \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} + \|\nabla \mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} \\ &= \gamma^{-1} \|\nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} + \|\nabla \mathbf{w}_h\|_{\mathbf{L}^2(\Omega)}. \end{aligned}$$

Step 5. We estimate T_2 . Let $\ell, k \in \{1, 2, 3\}$. Elementwise, it holds that

$$\partial_\ell \mathbf{T}_n \partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] = \mathbf{T}_n \partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] \stackrel{(53d)}{=} \mathbf{0}$$

as well as $\partial_\ell \partial_k \mathbf{w}_h = \mathbf{0}$. Together with the product rule, this yields that

$$\begin{aligned} \partial_\ell \partial_k [\mathbf{q}[\boldsymbol{\mu}_h] \mathbf{w}_h] &\stackrel{(55)}{=} \partial_k [\mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_\ell \mathbf{w}_h - \partial_\ell \mathbf{T}_n \partial_k [\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \mathbf{w}_h - \partial_k [\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_\ell \mathbf{w}_h \\ &\quad + \partial_\ell \mathbf{T}_n [\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_k \mathbf{w}_h - \partial_\ell [\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_k \mathbf{w}_h \\ &= \mathbf{T}_n \partial_k [\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_\ell \mathbf{w}_h - \mathbf{T}_n \partial_\ell \partial_k [\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \mathbf{w}_h - \mathbf{T}_n \partial_k [\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_\ell \mathbf{w}_h \\ &\quad + \mathbf{T}_n \partial_\ell [\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_k \mathbf{w}_h - \mathbf{T}_n \partial_\ell [\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h]] \partial_k \mathbf{w}_h. \end{aligned}$$

Recall that $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$. With (53) and $1 \leq 2/\gamma$, the latter equation yields that

$$\begin{aligned} T_2 &\stackrel{(54)}{\lesssim} \gamma^{-1} \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} + \gamma^{-2} \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h\|_{L^\infty(\Omega)}^2 \|\mathbf{w}_h\|_{L^2(\Omega)} \\ &= \gamma^{-1} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} + \gamma^{-2} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)}^2 \|\mathbf{w}_h\|_{L^2(\Omega)}. \end{aligned}$$

Step 6. We combine **Step 3–Step 5**. For all $\mathbf{w}_h \in (\mathcal{S}_h)^2$, this yields that

$$\begin{aligned} \|\nabla [\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h]\|_{L^2(\Omega)} &\lesssim \gamma^{-1} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} \\ &\quad + h \gamma^{-1} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} + h \gamma^{-2} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)}^2 \|\mathbf{w}_h\|_{L^2(\Omega)}. \end{aligned}$$

With an inverse estimate and $1 \leq 2/\gamma$, the latter equation yields for all $\mathbf{w}_h \in (\mathcal{S}_h)^2$ that

$$\|\nabla [\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h]\|_{L^2(\Omega)} \lesssim \gamma^{-2} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \|\nabla \mathbf{w}_h\|_{L^2(\Omega)}. \quad (56)$$

This concludes the proof of (i).

For the proof of (ii), let $\tilde{\mathcal{I}}_h$ be the nodal interpolant in 2D. With $\tilde{\mathcal{I}}_h$ instead of \mathcal{I}_h and $\mathbf{q}[\boldsymbol{\mu}_h]^T : \Omega \rightarrow \mathbb{R}^{2 \times 3}$ instead of $\mathbf{q}[\boldsymbol{\mu}_h]$, the proof of (ii) follows the lines of **Step 1–Step 5**.

For the proof of (iii), let $\mathbf{v}_h \in \mathcal{S}_h$ and $\mathbf{w}_h := \mathbb{P}_h^T[\boldsymbol{\mu}_h] \mathbf{v}_h \in \mathcal{S}_h$. With Lemma 10 (iii) and Lemma 11 (ii), we get that

$$\begin{aligned} \|\nabla [\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{v}_h]\|_{L^2(\Omega)} &= \|\nabla [(\mathbb{P}_h[\boldsymbol{\mu}_h] \circ \mathbb{P}_h^T[\boldsymbol{\mu}_h]) \mathbf{v}_h]\|_{L^2(\Omega)} \lesssim \|\nabla [\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{w}_h]\|_{L^2(\Omega)} \\ &\stackrel{(i)}{\lesssim} \gamma^{-2} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\mathbb{P}_h^T[\boldsymbol{\mu}_h] \mathbf{v}_h\|_{L^2(\Omega)} + \|\nabla [\mathbb{P}_h^T[\boldsymbol{\mu}_h] \mathbf{v}_h]\|_{L^2(\Omega)} \\ &\lesssim \gamma^{-2} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\mathbf{v}_h\|_{L^2(\Omega)} + \|\nabla [\mathbb{P}_h^T[\boldsymbol{\mu}_h] \mathbf{v}_h]\|_{L^2(\Omega)} \\ &\stackrel{(ii)}{\lesssim} \gamma^{-2} \|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\mathbf{v}_h\|_{L^2(\Omega)} + \|\nabla \mathbf{v}_h\|_{L^2(\Omega)}. \end{aligned}$$

This proves (iii) and concludes the proof. \square

In the following lemma, we prove a discrete \mathbf{H}^1 -continuity of the mapping $\mathbb{P}_h(\cdot)$ from (46). Unlike Lemma 10 and Lemma 11, the following lemma builds on the explicit definition of the Householder matrices (17).

Lemma 13. Let $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h \in \mathcal{M}_h$ with $1 + (\mathbf{T}_n \boldsymbol{\mu}_h(z))_3 \geq \gamma > 0$ and $1 + (\mathbf{T}_n \boldsymbol{\nu}_h(z))_3 \geq \gamma > 0$ for all $z \in \mathcal{N}_h$. Then, there exists $C > 1$, which depends only on C_{mesh} , such that

$$\begin{aligned} & \|\nabla((\mathbb{P}_h[\boldsymbol{\mu}_h] - \mathbb{P}_h[\boldsymbol{\nu}_h]) \mathbf{w}_h)\|_{\mathbf{L}^2(\Omega)} \\ & \leq C \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} + C \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{\mathbf{L}^2(\Omega)} \\ & \quad + C \gamma^{-3} (\|\nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)}) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} \|\mathbf{w}_h\|_{\mathbf{L}^2(\Omega)}, \end{aligned} \quad (57)$$

for all $\mathbf{w}_h \in (\mathcal{S}_h)^2$.

Proof. We split the proof into the following six steps.

Step 1. With the assumption $1 + (\mathbf{T}_n \boldsymbol{\mu}_h(z))_3 \geq \gamma > 0$ and $1 + (\mathbf{T}_n \boldsymbol{\nu}_h(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, we use the definitions (50) of $\mathbf{R}_1(\cdot)$ and $\mathbf{R}_2(\cdot)$ and interpret

$$\mathbf{q}[\boldsymbol{\mu}_h] := \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h], \quad (58a)$$

$$\mathbf{q}[\boldsymbol{\nu}_h] := \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h] - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h] \quad (58b)$$

as functions $\mathbf{q}[\boldsymbol{\mu}_h], \mathbf{q}[\boldsymbol{\nu}_h] : \Omega \rightarrow \mathbb{R}^{3 \times 2}$. With \mathcal{I}_h being the vector-valued nodal interpolant onto \mathcal{S}_h , recall from (52) that

$$(\mathbb{P}_h[\boldsymbol{\mu}_h] - \mathbb{P}_h[\boldsymbol{\nu}_h]) \mathbf{w}_h = \mathcal{I}_h((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h) \quad \text{for all } \mathbf{w}_h \in (\mathcal{S}_h)^2. \quad (59)$$

Step 2. Recall that $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$. With the definition (50) of $\mathbf{R}_1(\cdot)$, we get that

$$\|\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h]\|_{\mathbf{L}^\infty(\Omega)} \lesssim \|\mathbf{T}_n \boldsymbol{\mu}_h - \mathbf{T}_n \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} = \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)}. \quad (60a)$$

With the product rule, we further get for all $\ell, k \in \{1, 2, 3\}$ that

$$\begin{aligned} & \|\partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h]\|_{\mathbf{L}^\infty(\Omega)} \lesssim \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h - \nabla \mathbf{T}_n \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} \\ & = \|\mathbf{T}_n \nabla \boldsymbol{\mu}_h - \mathbf{T}_n \nabla \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} = \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} \end{aligned} \quad (60b)$$

as well as

$$\partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] = \mathbf{0} = \partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h]. \quad (60c)$$

Moreover, define

$$\begin{aligned} \sigma(\mathbf{T}_n \boldsymbol{\mu}_h, \mathbf{T}_n \boldsymbol{\nu}_h) & := \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} + \|\nabla \mathbf{T}_n \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} \\ & = \|\mathbf{T}_n \nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} + \|\mathbf{T}_n \nabla \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} \\ & = \|\nabla \boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} = \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h). \end{aligned} \quad (61a)$$

Note that an inverse inequality yields that

$$h \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \lesssim \|\boldsymbol{\mu}_h\|_{\mathbf{L}^\infty(\Omega)} + \|\boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} = 2. \quad (61b)$$

Lemma A.2 and the definition (50) of $\mathbf{R}_2(\cdot)$ then yield that

$$\|\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h]\|_{\mathbf{L}^\infty(\Omega)} \lesssim \gamma^{-1} \|\mathbf{T}_n \boldsymbol{\mu}_h - \mathbf{T}_n \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)} = \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{\mathbf{L}^\infty(\Omega)}. \quad (62a)$$

For all $\ell, k \in \{1, 2, 3\}$, we further get that

$$\begin{aligned}
& \|\partial_k \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \partial_k \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h]\|_{L^\infty(\Omega)} \\
& \lesssim \gamma^{-2} \sigma(\mathbf{T}_n \boldsymbol{\mu}_h, \mathbf{T}_n \boldsymbol{\nu}_h) \|\mathbf{T}_n \boldsymbol{\mu}_h - \mathbf{T}_n \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} + \gamma^{-1} \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h - \nabla \mathbf{T}_n \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \\
& \stackrel{(61)}{=} \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\mathbf{T}_n \boldsymbol{\mu}_h - \mathbf{T}_n \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} + \gamma^{-1} \|\mathbf{T}_n \nabla \boldsymbol{\mu}_h - \mathbf{T}_n \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \\
& = \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} + \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}
\end{aligned} \tag{62b}$$

as well as

$$\begin{aligned}
& \|\partial_\ell \partial_k \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \partial_\ell \partial_k \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h]\|_{L^\infty(\Omega)} \\
& \lesssim \gamma^{-3} \sigma(\mathbf{T}_n \boldsymbol{\mu}_h, \mathbf{T}_n \boldsymbol{\nu}_h)^2 \|\mathbf{T}_n \boldsymbol{\mu}_h - \mathbf{T}_n \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \\
& \quad + \gamma^{-2} \sigma(\mathbf{T}_n \boldsymbol{\mu}_h, \mathbf{T}_n \boldsymbol{\nu}_h) \|\nabla \mathbf{T}_n \boldsymbol{\mu}_h - \nabla \mathbf{T}_n \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \\
& \stackrel{(61)}{=} \gamma^{-3} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h)^2 \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}.
\end{aligned} \tag{62c}$$

Step 3. Let $\mathbf{w}_h \in (\mathcal{S}_h)^2$. Standard estimates for the nodal interpolant \mathcal{I}_h yield that

$$\begin{aligned}
& \|\nabla [(\mathbb{P}_h[\boldsymbol{\mu}_h] - \mathbb{P}_h[\boldsymbol{\nu}_h]) \mathbf{w}_h]\|_{L^2(\Omega)} \stackrel{(59)}{=} \|\nabla \mathcal{I}_h((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h)\|_{L^2(\Omega)} \\
& \leq \|\nabla((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h)\|_{L^2(\Omega)} + \|\nabla(\mathbf{1} - \mathcal{I}_h)((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h)\|_{L^2(\Omega)} \\
& \lesssim \|\nabla((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h)\|_{L^2(\Omega)} + h \left(\sum_{K \in \mathcal{T}_h} \|D^2((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h)\|_{L^2(K)}^2 \right)^{\frac{1}{2}} \\
& =: T_1 + h T_2.
\end{aligned}$$

Step 4. We estimate T_1 . Let $k \in \{1, 2, 3\}$. With the product rule, we get that

$$\begin{aligned}
& \partial_k((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h) = \partial_k((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h])) \mathbf{w}_h + (\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \partial_k \mathbf{w}_h \\
& \stackrel{(58)}{=} \partial_k((\mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h])) \mathbf{w}_h - \partial_k((\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h])) \mathbf{w}_h \\
& \quad + (\mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h]) \partial_k \mathbf{w}_h - (\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h]) \partial_k \mathbf{w}_h \\
& = \mathbf{T}_n \partial_k((\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h])) \mathbf{w}_h - \mathbf{T}_n \partial_k((\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h])) \mathbf{w}_h \\
& \quad + \mathbf{T}_n (\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h]) \partial_k \mathbf{w}_h - \mathbf{T}_n (\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h]) \partial_k \mathbf{w}_h.
\end{aligned}$$

Recall that $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$. With the estimates from **Step 2**, we further get that

$$\begin{aligned}
& \|\partial_k((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h)\|_{L^2(\Omega)} \\
& \leq \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} \\
& \quad + \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_k \mathbf{w}_h\|_{L^2(\Omega)} \\
& \quad + \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_k \mathbf{w}_h\|_{L^2(\Omega)}.
\end{aligned}$$

With $1 \leq 2/\gamma$, we arrive at

$$\begin{aligned}
T_1 & \lesssim \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} \\
& \quad + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)}.
\end{aligned}$$

Step 5. We estimate T_2 . To this end, let $\ell, k \in \{1, 2, 3\}$. Note that the second derivative of the piecewise affine function \mathbf{w}_h vanishes on each element $K \in \mathcal{T}_h$. Moreover, recall from (60c), that elementwise $\partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] = \mathbf{0} = \partial_\ell \partial_k \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h]$. The product rule yields elementwise that

$$\begin{aligned}
& \partial_\ell \partial_k \left((\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h \right) \\
&= \partial_\ell \partial_k \left(\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h] \right) \mathbf{w}_h + \partial_k \left(\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h] \right) \partial_\ell \mathbf{w}_h + \partial_\ell \left(\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h] \right) \partial_k \mathbf{w}_h \\
&\stackrel{(58)}{=} \partial_\ell \partial_k \left(\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \mathbf{w}_h + \partial_k \left(\mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_\ell \mathbf{w}_h \\
&\quad - \partial_k \left(\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_\ell \mathbf{w}_h + \partial_\ell \left(\mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_k \mathbf{w}_h \\
&\quad - \partial_\ell \left(\mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{T}_n \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_k \mathbf{w}_h \\
&= \mathbf{T}_n \partial_\ell \partial_k \left(\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \mathbf{w}_h + \mathbf{T}_n \partial_k \left(\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_\ell \mathbf{w}_h \\
&\quad - \mathbf{T}_n \partial_k \left(\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_\ell \mathbf{w}_h + \mathbf{T}_n \partial_\ell \left(\mathbf{R}_1[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_1[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_k \mathbf{w}_h \\
&\quad - \mathbf{T}_n \partial_\ell \left(\mathbf{R}_2[\mathbf{T}_n \boldsymbol{\mu}_h] - \mathbf{R}_2[\mathbf{T}_n \boldsymbol{\nu}_h] \right) \partial_k \mathbf{w}_h.
\end{aligned}$$

With the estimates from **Step 2** and since $\mathbf{T}_n = \mathbf{T}_n^{-1} = \mathbf{T}_n^T$, we get for all $K \in \mathcal{T}_h$ that

$$\begin{aligned}
& \|\partial_\ell \partial_k [(\mathbf{q}[\boldsymbol{\mu}_h] - \mathbf{q}[\boldsymbol{\nu}_h]) \mathbf{w}_h]\|_{L^2(K)} \\
&\lesssim \gamma^{-3} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h)^2 \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(K)} + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(K)} \\
&\quad + \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_\ell \mathbf{w}_h\|_{L^2(K)} + \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_\ell \mathbf{w}_h\|_{L^2(K)} \\
&\quad + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_\ell \mathbf{w}_h\|_{L^2(K)} + \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_k \mathbf{w}_h\|_{L^2(K)} \\
&\quad + \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_k \mathbf{w}_h\|_{L^2(K)} + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\partial_k \mathbf{w}_h\|_{L^2(K)}.
\end{aligned}$$

With $1 \leq 2/\gamma$, the latter estimate simplifies to

$$\begin{aligned}
T_2 &\lesssim \gamma^{-3} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h)^2 \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} \\
&\quad + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(K)} + \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)}.
\end{aligned}$$

Step 6. We combine **Step 3–Step 5**. An inverse estimate and $1 \leq 2/\gamma$ imply that

$$\begin{aligned}
& \|\nabla [(\mathbb{P}_h[\boldsymbol{\mu}_h] - \mathbb{P}_h[\boldsymbol{\nu}_h]) \mathbf{w}_h]\|_{L^2(\Omega)} \leq T_1 + hT_2 \\
&\lesssim \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} \\
&\quad + \gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + h\gamma^{-3} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h)^2 \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} \\
&\quad + h\gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} + h\gamma^{-2} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} \\
&\quad + h\gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} \\
&\stackrel{(61)}{\lesssim} \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)} + \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \mathbf{w}_h\|_{L^2(\Omega)} \\
&\quad + \gamma^{-3} \sigma(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbf{w}_h\|_{L^2(\Omega)}.
\end{aligned}$$

This concludes the proof. \square

6.2. Energy norms. From (24), recall for given $\boldsymbol{\mu}_h \in \mathcal{M}_h$ the energy-scalar product

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle_{\boldsymbol{\mu}_h} := \mathbf{x} \cdot (\mathbf{P}_Q[\boldsymbol{\mu}_h])^{-1} \mathbf{y} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}, \quad (63)$$

with the corresponding energy norm $\|\cdot\|_{\boldsymbol{\mu}_h}$. With the definitions of the matrices from Section 2.4 as well as the definition (46b) of $\tilde{\mathbb{P}}_h(\cdot)$, it follows that

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle_{\boldsymbol{\mu}_h} = \alpha_P \langle \tilde{\mathbb{P}}[\boldsymbol{\mu}_h] \mathbf{x}, \tilde{\mathbb{P}}[\boldsymbol{\mu}_h] \mathbf{y} \rangle_\Omega + \beta(k)k \langle \nabla \tilde{\mathbb{P}}[\boldsymbol{\mu}_h] \mathbf{x}, \nabla \tilde{\mathbb{P}}[\boldsymbol{\mu}_h] \mathbf{y} \rangle_\Omega. \quad (64)$$

This section collects equivalence results for varying arguments $\boldsymbol{\mu}_h$ in $\langle\langle \cdot, \cdot \rangle\rangle_{\boldsymbol{\mu}_h}$ and $\|\cdot\|_{\boldsymbol{\mu}_h}$.

Lemma 14. *Let $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h \in \mathcal{M}_h$ and*

$$\tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2}) := \left(1 + \frac{\beta(k)k}{\alpha_P h^2} \max_{z \in \mathcal{N}_h} \|\mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z)] - \mathbf{H}[\mathbf{T}_n \boldsymbol{\nu}_h(z)]\|^2 \right)^{1/2} \geq 1. \quad (65a)$$

Then, $\tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2}) = \tilde{\kappa}(\boldsymbol{\nu}_h, \boldsymbol{\mu}_h, h^{-2})$ and there exists a constant $C \geq 1$, which depends only on C_{mesh} , such that

$$C^{-1} \tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2})^{-1} \|\mathbf{x}\|_{\boldsymbol{\nu}_h} \leq \|\mathbf{x}\|_{\boldsymbol{\mu}_h} \leq C \tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2}) \|\mathbf{x}\|_{\boldsymbol{\nu}_h} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}. \quad (65b)$$

Proof. Let $\mathbf{x} \in \mathbb{R}^{2N}$. Since the symmetry $\tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2}) = \tilde{\kappa}(\boldsymbol{\nu}_h, \boldsymbol{\mu}_h, h^{-2})$ is obvious, we only have to show that $\|\mathbf{x}\|_{\boldsymbol{\mu}_h} \lesssim \tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2}) \|\mathbf{x}\|_{\boldsymbol{\nu}_h}$. To this end, Lemma 11 (v) and an inverse estimate yield that

$$\begin{aligned} \|\mathbf{x}\|_{\boldsymbol{\mu}_h}^2 &\stackrel{(64)}{=} \alpha_P \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 + \beta(k)k \|\nabla \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 \\ &\lesssim \alpha_P \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 + \beta(k)k \|\nabla \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 + \beta(k)k \|\nabla \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} - \nabla \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 \\ &\lesssim \|\mathbf{x}\|_{\boldsymbol{\nu}_h}^2 + \beta(k)kh^{-2} \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} - \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2. \end{aligned}$$

With Lemma 11 (iv) and (vi), we estimate the last term by

$$\begin{aligned} \beta(k)kh^{-2} \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} - \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 &\lesssim \beta(k)kh |\mathbf{x}|^2 \max_{z \in \mathcal{N}_h} \|\mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z)] - \mathbf{H}[\mathbf{T}_n \boldsymbol{\nu}_h(z)]\|^2 \\ &\simeq \beta(k)kh^{-2} \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 \max_{z \in \mathcal{N}_h} \|\mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z)] - \mathbf{H}[\mathbf{T}_n \boldsymbol{\nu}_h(z)]\|^2 \\ &\stackrel{(64)}{\lesssim} \beta(k)\alpha_P^{-1}kh^{-2} \|\mathbf{x}\|_{\boldsymbol{\nu}_h}^2 \max_{z \in \mathcal{N}_h} \|\mathbf{H}[\mathbf{T}_n \boldsymbol{\mu}_h(z)] - \mathbf{H}[\mathbf{T}_n \boldsymbol{\nu}_h(z)]\|^2. \end{aligned}$$

This proves $\|\mathbf{x}\|_{\boldsymbol{\mu}_h} \lesssim \tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2}) \|\mathbf{x}\|_{\boldsymbol{\nu}_h}$ and hence concludes the proof. \square

For certain $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h \in \mathcal{M}_h$, the norm equivalence $\|\cdot\|_{\boldsymbol{\mu}_h} \simeq \|\cdot\|_{\boldsymbol{\nu}_h}$ holds independently of the mesh-size h .

Lemma 15. *Let $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h \in \mathcal{M}_h$ with $1 + (\mathbf{T}_n \boldsymbol{\mu}_h(z))_3 \geq \gamma > 0$ and $1 + (\mathbf{T}_n \boldsymbol{\nu}_h(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. Let*

$$\begin{aligned} \kappa(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) &:= \left[1 + \gamma^{-2} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 + \frac{\beta(k)k}{\alpha_P \gamma^2} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 \right. \\ &\quad \left. + \frac{\beta(k)k}{\alpha_P \gamma^6} (\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)}^2 + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 \right]^{1/2} \geq 1. \end{aligned} \quad (66)$$

Then, $\kappa(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) = \kappa(\boldsymbol{\nu}_h, \boldsymbol{\mu}_h)$ and there exists $C \geq 1$ depending only on C_{mesh} such that

$$C^{-1} \kappa(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h)^{-1} \|\mathbf{x}\|_{\boldsymbol{\nu}_h} \leq \|\mathbf{x}\|_{\boldsymbol{\mu}_h} \leq C \kappa(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\mathbf{x}\|_{\boldsymbol{\nu}_h} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}.$$

Proof. Let $\mathbf{x} \in \mathbb{R}^{2N}$. Since the symmetry $\kappa(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) = \kappa(\boldsymbol{\nu}_h, \boldsymbol{\mu}_h)$ is obvious, we only have to show that $\|\mathbf{x}\|_{\boldsymbol{\mu}_h} \lesssim \kappa(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h) \|\mathbf{x}\|_{\boldsymbol{\nu}_h}$. With Lemma 10, we get that

$$\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] = \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] + (\mathbb{P}_h[\boldsymbol{\mu}_h] - \mathbb{P}_h[\boldsymbol{\nu}_h]) \circ \mathbb{P}_h^T[\boldsymbol{\nu}_h] \circ \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h]. \quad (67a)$$

With Lemma 13, Lemma 11 (ii), and Lemma 12 (ii), we get that

$$\begin{aligned} & \|\nabla(\mathbb{P}_h[\boldsymbol{\mu}_h] - \mathbb{P}_h[\boldsymbol{\nu}_h]) \circ \mathbb{P}_h^T[\boldsymbol{\nu}_h] \circ \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)} \\ & \stackrel{(57)}{\lesssim} \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbb{P}_h^T[\boldsymbol{\nu}_h] \circ \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)} \\ & \quad + \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla[\mathbb{P}_h^T[\boldsymbol{\nu}_h] \circ \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}]\|_{L^2(\Omega)} \\ & \quad + \gamma^{-3} (\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\mathbb{P}_h^T[\boldsymbol{\nu}_h] \circ \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)} \\ & \lesssim \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)} + \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\nabla \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)} \\ & \quad + \gamma^{-3} (\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}. \end{aligned} \quad (67b)$$

Recalling $\|\mathbb{P}_h[\boldsymbol{\mu}_h] \mathbf{x}\|_{L^2(\Omega)} \simeq \|\mathbb{P}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}$ from Lemma 11 (v), we obtain that

$$\begin{aligned} \|\mathbf{x}\|_{\boldsymbol{\mu}_h}^2 & \stackrel{(67)}{\lesssim} \alpha_{\mathbf{P}} \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 + \beta(k)k \left[\gamma^{-2} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 \right. \\ & \quad \left. + \gamma^{-6} (\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)}^2 + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 \right] \|\tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 \\ & \quad + \beta(k)k \left[1 + \gamma^{-2} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 \right] \|\nabla \tilde{\mathbb{P}}_h[\boldsymbol{\nu}_h] \mathbf{x}\|_{L^2(\Omega)}^2 \\ & \stackrel{(67)}{\leq} \left[1 + \gamma^{-2} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^2} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 \right. \\ & \quad \left. + \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^6} (\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)}^2 + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}^2 \right] \|\mathbf{x}\|_{\boldsymbol{\nu}_h}^2. \end{aligned}$$

This proves $\|\mathbf{x}\|_{\boldsymbol{\mu}_h} \lesssim \kappa(\boldsymbol{\nu}_h, \boldsymbol{\mu}_h) \|\mathbf{x}\|_{\boldsymbol{\nu}_h}$ and concludes the proof. \square

Lemma 16. Let $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h \in \mathcal{M}_h$. There exists a constant $C > 1$, which depends only on $C_{\text{mesh}} > 0$, such that the following two assertions (i)–(ii) hold true:

(i) With $\tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2})$ from (65a), it holds that, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}$,

$$\mathbf{x} \cdot \mathbf{A}_{\mathbf{Q}}[\boldsymbol{\mu}_h] \mathbf{x} \geq C^{-1} \frac{\alpha}{\alpha_{\mathbf{P}}} \tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2})^{-2} \|\mathbf{x}\|_{\boldsymbol{\nu}_h}^2, \quad \text{and} \quad (68a)$$

$$\mathbf{x} \cdot \mathbf{A}_{\mathbf{Q}}[\boldsymbol{\mu}_h] \mathbf{y} \leq C \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right) \tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2})^2 \|\mathbf{x}\|_{\boldsymbol{\nu}_h} \|\mathbf{y}\|_{\boldsymbol{\nu}_h}. \quad (68b)$$

(ii) If, additionally, $1 + (\mathbf{T}_n \boldsymbol{\mu}_h(z))_3 \geq \gamma > 0$ and $1 + (\mathbf{T}_n \boldsymbol{\nu}_h(z))_3 \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, the statement of (i) holds with $\kappa(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h)$ from (66) instead of $\tilde{\kappa}(\boldsymbol{\mu}_h, \boldsymbol{\nu}_h, h^{-2})$ from (65a). In particular, the estimate then is independent of the mesh-size h .

Proof. First, we prove (i). Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}$. Recall $\mathbf{A}_Q[\boldsymbol{\mu}_h]$ from (21) as well as $\mathbf{A}_k[\boldsymbol{\mu}_h]$, $\mathbf{M}_k[\boldsymbol{\mu}_h]$, \mathbf{L} , $\mathbf{S}[\boldsymbol{\mu}_h]$ from Section 2.4. Since $\mathbf{S}[\boldsymbol{\mu}_h]$ is skew-symmetric, it holds that

$$\begin{aligned} \mathbf{x} \cdot \mathbf{A}_Q[\boldsymbol{\mu}_h] \mathbf{x} &\stackrel{(21)}{=} \alpha \mathbf{Q}[\boldsymbol{\mu}_h] \mathbf{x} \cdot \mathbf{M}_k[\mathbf{m}_h^n] \mathbf{Q}[\boldsymbol{\mu}_h] \mathbf{x} + \beta(k)k \mathbf{Q}[\boldsymbol{\mu}_h] \mathbf{x} \cdot \mathbf{L} \mathbf{Q}[\boldsymbol{\mu}_h] \mathbf{x} \\ &\stackrel{(46a)}{=} \langle W_k(\lambda_h^n) \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}, \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} \rangle_\Omega + \beta(k)k \langle \nabla \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}, \nabla \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} \rangle_\Omega \\ &\stackrel{(8)}{\geq} \frac{\alpha}{2} \langle \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}, \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} \rangle_\Omega + \beta(k)k \langle \nabla \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}, \nabla \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x} \rangle_\Omega \stackrel{(64)}{\geq} \frac{\alpha}{2\alpha_{\mathbf{P}}} \|\mathbf{x}\|_{\boldsymbol{\mu}_h}^2. \end{aligned}$$

With the norm equivalence result from Lemma 14, we replace $\|\mathbf{x}\|_{\boldsymbol{\mu}_h}$ with $\|\mathbf{x}\|_{\boldsymbol{\nu}_h}$ and prove (68a). Similarly, we obtain that

$$\begin{aligned} \mathbf{x} \cdot \mathbf{A}_Q[\boldsymbol{\mu}_h] \mathbf{y} &\stackrel{(64)}{=} \langle \langle \mathbf{x}, \mathbf{y} \rangle \rangle_{\boldsymbol{\mu}_h} + \langle (W_k(\lambda_h^n) - \alpha_{\mathbf{P}}) \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}, \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{y} \rangle_\Omega \\ &\quad + \langle \boldsymbol{\mu}_h \times \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}, \tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{y} \rangle_\Omega \\ &\leq \|\mathbf{x}\|_{\boldsymbol{\mu}_h} \|\mathbf{y}\|_{\boldsymbol{\mu}_h} + (1 + \|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}) \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{x}\|_{L^2(\Omega)} \|\tilde{\mathbb{P}}_h[\boldsymbol{\mu}_h] \mathbf{y}\|_{L^2(\Omega)} \\ &\stackrel{(64)}{\leq} \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right) \|\mathbf{x}\|_{\boldsymbol{\mu}_h} \|\mathbf{y}\|_{\boldsymbol{\mu}_h}. \end{aligned}$$

Again, with the norm equivalence result from Lemma 14, we prove (68b). This concludes the proof of (i). The proof of (ii) follows the same lines but employs Lemma 15 instead Lemma 14. Altogether, this concludes the proof. \square

6.3. Proof of Theorem 3. Since $\mathbf{A}_k[\mathbf{m}_h^n]$ is positive definite and $\mathbf{Q}[\mathbf{m}_h^n]$ has orthonormal columns, the system matrix in (19) is also positive definite. Let $\mathbf{x} \in \mathbb{R}^{2N}$ be the unique solution of (19). Then, it holds that

$$\mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{A}_k[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \cdot \mathbf{y} \stackrel{(19)}{=} \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{b}[\mathbf{m}_h^n] \cdot \mathbf{y} \quad \text{for all } \mathbf{y} \in \mathbb{R}^{2N}. \quad (69)$$

We denote the bilinear form on the left-hand side of (9b) by $A_h(\cdot, \cdot)$ and the linear functional on the right-hand side of (9b) by $R(\cdot)$. The definition of $\mathbf{A}_k[\mathbf{m}_h^n]$ in Section 2.4 then yields that

$$\begin{aligned} A_h(\tilde{\mathbb{P}}_h[\mathbf{m}_h^n] \mathbf{x}, \tilde{\mathbb{P}}_h[\mathbf{m}_h^n] \mathbf{y}) &= \mathbf{A}_k[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \cdot \mathbf{Q}[\mathbf{m}_h^n] \mathbf{y} = \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{A}_k[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \cdot \mathbf{y} \\ &\stackrel{(69)}{=} \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{b}[\mathbf{m}_h^n] \cdot \mathbf{y} = \mathbf{b}[\mathbf{m}_h^n] \cdot \mathbf{Q}[\mathbf{m}_h^n] \mathbf{y} = R(\tilde{\mathbb{P}}_h[\mathbf{m}_h^n] \mathbf{y}) \quad \text{for all } \mathbf{y} \in \mathbb{R}^{2N}. \end{aligned}$$

With Lemma 10 (v), $\tilde{\mathbb{P}}_h[\mathbf{m}_h^n]$ is an isomorphism from \mathbb{R}^{2N} to $\mathcal{K}_h[\mathbf{m}_h^n]$. Consequently, the function $\tilde{\mathbb{P}}_h[\mathbf{m}_h^n] \mathbf{x} \in \mathcal{K}_h[\mathbf{m}_h^n]$ is a solution to (9b). The representation formula (20) follows from (46b). This concludes the proof. \square

6.4. Proof of Theorem 5. First, we prove (i). For a non-symmetric but positive definite system matrix, the fields-of-value analysis for the preconditioned GMRES algorithm (see, e.g., [Sta97, Theorem 3.2]) yields that

$$\|\mathbf{r}^{(\ell)}\|_{\boldsymbol{\mu}_h} \leq \left(1 - \gamma^{(1)} \gamma^{(2)} \right)^{\ell/2} \|\mathbf{r}^{(0)}\|_{\boldsymbol{\mu}_h}, \quad (70a)$$

where

$$\gamma^{(1)} := \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \mathbf{A}_Q[\mathbf{m}_h^n] \mathbf{x}}{\mathbf{x} \cdot (\mathbf{P}_Q[\boldsymbol{\mu}_h])^{-1} \mathbf{x}} > 0, \quad (70b)$$

$$\gamma^{(2)} := \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot (\mathbf{A}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}}{\mathbf{x} \cdot \mathbf{P}_Q[\boldsymbol{\mu}_h] \mathbf{x}} > 0. \quad (70c)$$

To estimate $\gamma^{(1)}$ and $\gamma^{(2)}$ from below, recall $\tilde{\kappa}(\mathbf{m}_h^n, \boldsymbol{\mu}_h, h^{-2})$ from (65a) and exploit Lemma 16 (i). This yields that

$$\gamma^{(1)} \stackrel{(65)}{\gtrsim} \frac{\alpha}{\alpha_{\mathbf{P}}} \tilde{\kappa}(\mathbf{m}_h^n, \boldsymbol{\mu}_h, h^{-2})^{-2} \stackrel{(26)}{=} \frac{\alpha}{\alpha_{\mathbf{P}}} \mathbb{F}^{-1}.$$

With Lemma 16 (i), the matrices $\mathbf{B} := \mathbf{A}_Q[\mathbf{m}_h^n]$ and $\mathbf{B}_0 := (\mathbf{P}_Q[\boldsymbol{\mu}_h])^{-1}$ satisfy the setting of Lemma A.3 with

$$c_1 \simeq \frac{\alpha}{\alpha_{\mathbf{P}}} \tilde{\kappa}(\mathbf{m}_h^n, \boldsymbol{\mu}_h, h^{-2})^{-2}, \quad \text{and} \quad c_2 \simeq \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right) \tilde{\kappa}(\mathbf{m}_h^n, \boldsymbol{\mu}_h, h^{-2})^2.$$

Hence, Lemma A.3 yields that

$$\begin{aligned} \gamma^{(2)} &\stackrel{(70c)}{\gtrsim} \frac{\alpha}{\alpha_{\mathbf{P}}} \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right)^{-2} \tilde{\kappa}(\mathbf{m}_h^n, \boldsymbol{\mu}_h, h^{-2})^{-6} \\ &\stackrel{(26)}{=} \frac{\alpha}{\alpha_{\mathbf{P}}} \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right)^{-2} \mathbb{F}^{-3}. \end{aligned} \quad (71)$$

With (70), we conclude the proof of (i). The proof of (ii) then follows the same lines but exploits Lemma 16 (ii) instead of Lemma 16 (i). In the latter arguments, this replaces $\tilde{\kappa}(\mathbf{m}_h^n, \boldsymbol{\mu}_h, h^{-2})$ by $\kappa(\mathbf{m}_h^n, \boldsymbol{\mu}_h)$. Altogether, this concludes the proof. \square

6.5. Proof of Theorem 7. In analogy to (70), the fields-of-value analysis for the preconditioned GMRES algorithm (see, e.g., [Sta97, Theorem 3.2]) yields that

$$\|\tilde{\mathbf{r}}^{(\ell)}\|'_{\mathbf{m}_h^n} \leq \left(1 - \tilde{\gamma}^{(1)} \tilde{\gamma}^{(2)} \right)^{\ell/2} \|\tilde{\mathbf{r}}^{(0)}\|'_{\mathbf{m}_h^n}, \quad (72a)$$

where

$$\tilde{\gamma}^{(1)} := \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \mathbf{A}_Q[\mathbf{m}_h^n] \mathbf{x}}{\mathbf{x} \cdot (\tilde{\mathbf{P}}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}} > 0, \quad (72b)$$

$$\tilde{\gamma}^{(2)} := \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot (\mathbf{A}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}}{\mathbf{x} \cdot \tilde{\mathbf{P}}_Q[\mathbf{m}_h^n] \mathbf{x}} > 0. \quad (72c)$$

Recall from (23), the definition of the theoretical preconditioner $\mathbf{P}[\mathbf{m}_h^n]$ and from (70) the corresponding definition of $\gamma^{(1)}$ and $\gamma^{(2)}$. We obtain that

$$\begin{aligned} \tilde{\gamma}^{(1)} &\stackrel{(72b)}{\geq} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot \mathbf{A}_Q[\mathbf{m}_h^n] \mathbf{x}}{\mathbf{x} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}}{\mathbf{x} \cdot (\tilde{\mathbf{P}}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}} \\ &\stackrel{(70b)}{=} \gamma^{(1)} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{\mathbf{0}\}} \frac{\mathbf{x} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}}{\mathbf{x} \cdot (\tilde{\mathbf{P}}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}} =: \gamma^{(1)} \delta^{(1)}, \end{aligned} \quad (73a)$$

as well as

$$\begin{aligned}\tilde{\gamma}^{(2)} &\stackrel{(72b)}{\geq} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{0\}} \frac{\mathbf{x} \cdot (\mathbf{A}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}}{\mathbf{x} \cdot \mathbf{P}_Q[\mathbf{m}_h^n] \mathbf{x}} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{0\}} \frac{\mathbf{x} \cdot \mathbf{P}_Q[\mathbf{m}_h^n] \mathbf{x}}{\mathbf{x} \cdot \tilde{\mathbf{P}}_Q[\mathbf{m}_h^n] \mathbf{x}} \\ &\stackrel{(70b)}{=} \gamma^{(2)} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{0\}} \frac{\mathbf{x} \cdot \mathbf{P}_Q[\mathbf{m}_h^n] \mathbf{x}}{\mathbf{x} \cdot \tilde{\mathbf{P}}_Q[\mathbf{m}_h^n] \mathbf{x}} =: \gamma^{(2)} \delta^{(2)}.\end{aligned}\quad (73b)$$

Here, we implicitly have $\mathbf{m}_h^n = \boldsymbol{\mu}_h$ in the definition (70) of $\gamma^{(1)}$ and $\gamma^{(2)}$. In particular, Lemma 16 holds with $\tilde{\kappa}(\mathbf{m}_h^n, \mathbf{m}_h^n, h^{-2}) = \kappa(\mathbf{m}_h^n, \mathbf{m}_h^n) = 1$. Following the lines of the proof of Theorem 5, this yields that

$$\gamma^{(1)} \gtrsim \frac{\alpha}{\alpha_{\mathbf{P}}}, \quad \text{and} \quad \gamma^{(2)} \stackrel{(71)}{\gtrsim} \frac{\alpha}{\alpha_{\mathbf{P}}} \left(1 + \frac{1}{\alpha_{\mathbf{P}}} + \frac{\|W_k - \alpha_{\mathbf{P}}\|_{L^\infty(\Omega)}}{\alpha_{\mathbf{P}}} \right)^{-2}. \quad (74)$$

Hence, in the following four steps, it remains to estimate $\delta^{(1)}$ and $\delta^{(2)}$ from below.

Step 1. We will use the fictitious space lemma (see [Nep91, GO95]) to derive

$$c_1 \mathbf{x} \cdot \mathbf{P}_Q[\mathbf{m}_h^n] \mathbf{x} \leq \mathbf{x} \cdot \tilde{\mathbf{P}}_Q[\mathbf{m}_h^n] \mathbf{x} \leq c_2 \mathbf{x} \cdot \mathbf{P}_Q[\mathbf{m}_h^n] \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}. \quad (75)$$

Here, the constants $c_1, c_2 > 0$ stem from the following two assumptions **(FS1)**–**(FS2)** of the fictitious space lemma:

(FS1) For all $\mathbf{x} \in \mathbb{R}^{2N}$, there exists $\mathbf{y} \in \mathbb{R}^{3N}$ with $\mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{y} = \mathbf{x}$ and

$$c_1 \mathbf{y} \cdot (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{y} \leq \mathbf{x} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x}. \quad (76a)$$

(FS2) For all $\mathbf{y} \in \mathbb{R}^{3N}$, it holds that

$$\mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{y} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{y} \leq c_2 \mathbf{y} \cdot (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{y}. \quad (76b)$$

With the assumptions **(FS1)**–**(FS2)**, the fictitious space lemma then implies that

$$\begin{aligned}c_1 \tilde{\mathbf{x}} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \tilde{\mathbf{x}} &\leq (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \tilde{\mathbf{x}} \cdot \mathbf{Q}[\mathbf{m}_h^n]^T (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L})^{-1} \mathbf{Q}[\mathbf{m}_h^n] (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \tilde{\mathbf{x}} \\ &\stackrel{(32)}{=} (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \tilde{\mathbf{x}} \cdot \tilde{\mathbf{P}}_Q[\mathbf{m}_h^n] (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \tilde{\mathbf{x}} \\ &\leq c_2 \tilde{\mathbf{x}} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \tilde{\mathbf{x}} \quad \text{for all } \tilde{\mathbf{x}} \in \mathbb{R}^{2N}.\end{aligned}$$

Then, with $\tilde{\mathbf{x}} := \mathbf{P}_Q[\mathbf{m}_h^n] \mathbf{x} \in \mathbb{R}^{2N}$ in the latter estimate, this verifies (75).

Step 2. We verify assumption **(FS1)** of the fictitious space lemma. To that end, let $\mathbf{x} \in \mathbb{R}^{2N}$ and set $\mathbf{y} := \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \in \mathbb{R}^{3N}$. Then, $\mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{y} = \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} = \mathbf{x}$ and

$$\begin{aligned}\mathbf{y} \cdot (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{y} &= \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \cdot (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \\ &= \mathbf{x} \cdot \mathbf{Q}[\mathbf{m}_h^n]^T (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{Q}[\mathbf{m}_h^n] \mathbf{x} \stackrel{(23)}{=} \mathbf{x} \cdot (\mathbf{P}_Q[\mathbf{m}_h^n])^{-1} \mathbf{x},\end{aligned}$$

i.e., assumption **(FS1)** holds with $c_1 = 1$.

Step 3. We verify assumption **(FS2)** of the fictitious space lemma. To that end, let $\mathbf{y} \in \mathbb{R}^{3N}$. Define $\mathbf{v}_h := \sum_{i=1}^{3N} \mathbf{y}_i \phi_i \in \mathcal{S}_h$. With Lemma 10 (ii), we obtain that

$$\begin{aligned}
& \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{y} \cdot (\mathbf{P}[\mathbf{m}_h^n])^{-1} \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{y} \\
& \stackrel{(23)}{=} \mathbf{y} \cdot \mathbf{Q}[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n]^T (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{Q}[\mathbf{m}_h^n] \mathbf{Q}[\mathbf{m}_h^n]^T \mathbf{y} \\
& \stackrel{(48b)}{=} \alpha_{\mathbf{P}} \|\tilde{\Pi}_h[\mathbf{m}_h^n] \mathbf{y}\|_{\mathbf{L}^2(\Omega)}^2 + \beta(k)k \|\nabla \tilde{\Pi}_h[\mathbf{m}_h^n] \mathbf{y}\|_{\mathbf{L}^2(\Omega)}^2 \\
& = \alpha_{\mathbf{P}} \|\Pi_h[\mathbf{m}_h^n] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 + \beta(k)k \|\nabla \Pi_h[\mathbf{m}_h^n] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2. \tag{77}
\end{aligned}$$

For the verification of **(FS2)** in (i), Lemma 11 (iii) and an inverse estimate yield that

$$\begin{aligned}
& \alpha_{\mathbf{P}} \|\Pi_h[\mathbf{m}_h^n] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 + \beta(k)k \|\nabla \Pi_h[\mathbf{m}_h^n] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 \\
& \lesssim \left(\alpha_{\mathbf{P}} + \frac{\beta(k)k}{h^2} \right) \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 \lesssim \left(1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} h^2} \right) \left[\alpha_{\mathbf{P}} \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 + \beta(k)k \|\nabla \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 \right] \\
& = \left(1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} h^2} \right) \mathbf{y} \cdot (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{y}. \tag{78a}
\end{aligned}$$

For the verification of **(FS2)** in (ii), we use the stronger assumption $1 + (\mathbf{T}_n \mathbf{m}_h^n)_3 \geq \gamma > 0$. Then, the definition of \mathbf{v}_h , Lemma 11 (iii) and Lemma 12 (iii) yield that

$$\begin{aligned}
& \alpha_{\mathbf{P}} \|\Pi_h[\mathbf{m}_h^n] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 + \beta(k)k \|\nabla \Pi_h[\mathbf{m}_h^n] \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 \\
& \lesssim \left(\alpha_{\mathbf{P}} + \frac{\beta(k)k}{\gamma^4} \|\nabla \mathbf{m}_h^n\|_{\mathbf{L}^\infty(\Omega)}^2 \right) \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 + \beta(k)k \|\nabla \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 \\
& \leq \left(1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^4} \|\nabla \mathbf{m}_h^n\|_{\mathbf{L}^\infty(\Omega)}^2 \right) \left[\alpha_{\mathbf{P}} \|\mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 + \beta(k)k \|\nabla \mathbf{v}_h\|_{\mathbf{L}^2(\Omega)}^2 \right] \\
& = \left(1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^4} \|\nabla \mathbf{m}_h^n\|_{\mathbf{L}^\infty(\Omega)}^2 \right) \mathbf{y} \cdot (\alpha_{\mathbf{P}} \mathbf{M} + \beta(k)k\mathbf{L}) \mathbf{y}. \tag{78b}
\end{aligned}$$

We combine (77)–(78) and obtain that **(FS2)** holds with

$$c_2 \lesssim \begin{cases} 1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} \gamma^4} \|\nabla \mathbf{m}_h^n\|_{\mathbf{L}^\infty(\Omega)}^2 & \text{if } 1 + (\mathbf{T}_n \mathbf{m}_h^n)_3 \geq \gamma > 0, \\ 1 + \frac{\beta(k)k}{\alpha_{\mathbf{P}} h^2} & \text{else.} \end{cases} \tag{79}$$

Step 4. With **Step 1–Step 3**, the matrices $\mathbf{B} := \tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n]$ and $\mathbf{B}_0 := \mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n]$ satisfy the assumptions of Lemma A.3 with $c_1 = 1$ and c_2 from (79). Hence, we get that

$$\frac{1}{c_2^2} \mathbf{x} \cdot (\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n])^{-1} \mathbf{x} \leq \mathbf{x} \cdot (\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n])^{-1} \mathbf{x} \leq \mathbf{x} \cdot (\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n])^{-1} \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N}. \tag{80}$$

From this and (75), we obtain that

$$\delta^{(1)} \stackrel{(73a)}{=} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{0\}} \frac{\mathbf{x} \cdot (\mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n])^{-1} \mathbf{x}}{\mathbf{x} \cdot (\tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n])^{-1} \mathbf{x}} \stackrel{(80)}{\geq} 1 \quad \text{and} \quad \delta^{(2)} \stackrel{(73b)}{=} \inf_{\mathbf{x} \in \mathbb{R}^{2N} \setminus \{0\}} \frac{\mathbf{x} \cdot \mathbf{P}_{\mathbf{Q}}[\mathbf{m}_h^n] \mathbf{x}}{\mathbf{x} \cdot \tilde{\mathbf{P}}_{\mathbf{Q}}[\mathbf{m}_h^n] \mathbf{x}} \stackrel{(75)}{\geq} \frac{1}{c_2}.$$

Together with the estimates for $\gamma^{(1)}$ and $\gamma^{(2)}$ from (74), this concludes the proof. \square

APPENDIX A. AUXILIARY RESULTS

Lemma A.1. For $\boldsymbol{\mu}_h = (\mu_1, \mu_2, \mu_3)^T \in \mathcal{M}_h$, there hold the following assertions (i)–(iii):

(i) For all $i, j \in \{1, 2\}$, it holds that

$$\left\| \frac{\mu_i \mu_j}{1 + \mu_3} \right\|_{L^\infty(\Omega)} \leq 2.$$

(ii) Let $1 + \mu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. For all $i, j \in \{1, 2\}$, it holds that

$$\frac{\mu_i \mu_j}{1 + \mu_3} \in W^{1,\infty}(\Omega) \quad \text{with} \quad \left\| \partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] \right\|_{L^\infty(\Omega)} \leq 3 \gamma^{-1} \|\partial_k \boldsymbol{\mu}_h\|_{L^\infty(\Omega)}$$

for all $k \in \{1, 2, 3\}$.

(iii) Let $1 + \mu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. For all $i, j \in \{1, 2\}$, it holds that

$$\frac{\mu_i \mu_j}{1 + \mu_3} \in W^{2,\infty}(K) \quad \text{with} \quad \left\| \partial_\ell \partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] \right\|_{L^\infty(K)} \leq 12 \gamma^{-2} \|\partial_k \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} \|\partial_\ell \boldsymbol{\mu}_h\|_{L^\infty(\Omega)},$$

for all elements $K \in \mathcal{T}_h$ and for all $\ell, k \in \{1, 2, 3\}$.

Proof. To prove (i), note that piecewise affine functions attain their maximal length at the nodes. Since $\boldsymbol{\mu}_h \in \mathcal{M}_h$, together with Young's inequality this yield that

$$\left| \frac{\mu_i \mu_j}{1 + \mu_3} \right| \leq \frac{\mu_1^2 + \mu_2^2}{1 + \mu_3} \leq \frac{1 - \mu_3^2}{1 + \mu_3} = \frac{(1 - \mu_3)(1 + \mu_3)}{1 + \mu_3} = (1 - \mu_3) \leq 2.$$

This proves (i). For the proof of (ii) and (iii), the product rule yields that

$$\partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] = \frac{(\partial_k \mu_i) \mu_j}{1 + \mu_3} + \frac{\mu_i (\partial_k \mu_j)}{1 + \mu_3} - \frac{\mu_i \mu_j (\partial_k \mu_3)}{(1 + \mu_3)^2}. \quad (81)$$

Moreover, we exploit that the second derivative of affine functions is zero. Elementwise, the product rule then yields that

$$\begin{aligned} \partial_\ell \partial_k \left[\frac{\mu_i \mu_j}{1 + \mu_3} \right] &\stackrel{(81)}{=} \frac{(\partial_k \mu_i) (\partial_\ell \mu_j)}{1 + \mu_3} - \frac{(\partial_k \mu_i) \mu_j (\partial_\ell \mu_3)}{(1 + \mu_3)^2} + \frac{(\partial_\ell \mu_i) (\partial_k \mu_j)}{1 + \mu_3} - \frac{\mu_i (\partial_k \mu_j) (\partial_\ell \mu_3)}{(1 + \mu_3)^2} \\ &\quad - \frac{(\partial_\ell \mu_i) \mu_j (\partial_k \mu_3)}{(1 + \mu_3)^2} - \frac{\mu_i (\partial_\ell \mu_j) (\partial_k \mu_3)}{(1 + \mu_3)^2} + \frac{2 \mu_i \mu_j (\partial_k \mu_3) (\partial_\ell \mu_3)}{(1 + \mu_3)^3}. \end{aligned} \quad (82)$$

Since $|\boldsymbol{\mu}_h| \leq 1$ and since $\boldsymbol{\mu}_h$ is piecewise affine, the assumption $1 + \mu_3(z) \geq \gamma > 0$ for all nodes implies that $1 + \mu_3 \geq \gamma > 0$ in Ω . Together with $|\boldsymbol{\mu}_h| \leq 1$ and (81), this proves (ii). For the proof of (iii), additionally note that $\gamma \leq 1 + \mu_3 \leq 1 + |\mu_3| \leq 2$ yields that $1 \leq 2/\gamma$. Together with $|\boldsymbol{\mu}_h| \leq 1$ and (82), this proves (iii). \square

Lemma A.2. For $\boldsymbol{\mu}_h = (\mu_1, \mu_2, \mu_3)^T, \boldsymbol{\nu}_h = (\nu_1, \nu_2, \nu_3)^T \in \mathcal{M}_h$, let $1 + \mu_3(z) \geq \gamma > 0$ and $1 + \nu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$. Then, there exists a constant $C > 0$ such that the following assertions (i)–(iii) hold true:

(i) For all $i, j \in \{1, 2\}$, it holds that

$$\left\| \frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} \right\|_{L^\infty(\Omega)} \leq C \gamma^{-1} \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}.$$

(ii) For all $i, j \in \{1, 2\}$ and all $k \in \{1, 2, 3\}$, it holds that

$$\left\| \partial_k \left(\frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} \right) \right\|_{L^\infty(\Omega)} \leq C \gamma^{-2} \left(\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \right) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \\ + C \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}.$$

(iii) For all $i, j \in \{1, 2\}$, all elements $K \in \mathcal{T}_h$ and all $\ell, k \in \{1, 2, 3\}$, it holds that

$$\left\| \partial_\ell \partial_k \left(\frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} \right) \right\|_{L^\infty(K)} \leq C \gamma^{-3} \left(\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \right)^2 \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \\ + C \gamma^{-2} \left(\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \right) \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}.$$

Proof. Throughout the proof, we write

$$P_{ij} := \frac{\mu_i \mu_j}{1 + \mu_3} \quad \text{and} \quad d_k := \mu_k - \nu_k \quad \text{for all } i, j \in \{1, 2\} \text{ and all } k \in \{1, 2, 3\}. \quad (83)$$

Recall that the properties of P_{ij} are discussed in Lemma A.1. Since $\boldsymbol{\mu}_h, \boldsymbol{\nu}_h$ are piecewise affine, we get that $|\boldsymbol{\mu}_h|, |\boldsymbol{\nu}_h| \leq 1$ on Ω . Moreover, since $1 + \mu_3(z) \geq \gamma > 0$ and $1 + \nu_3(z) \geq \gamma > 0$ for all nodes $z \in \mathcal{N}_h$, it follows that $1 + \mu_3 \geq \gamma > 0$ and $1 + \nu_3 \geq \gamma > 0$ on Ω . For the proof of (i), elementary computations show that

$$\frac{\mu_i \mu_j}{1 + \mu_3} - \frac{\nu_i \nu_j}{1 + \nu_3} = -P_{ij} \frac{d_3}{1 + \nu_3} + \frac{\mu_j d_i}{1 + \nu_3} + \frac{\nu_i d_j}{1 + \nu_3}. \quad (84)$$

Together with Lemma A.1 (i), this proves (i). For the proof of (ii), let $k \in \{1, 2, 3\}$. We differentiate the terms in (84) separately and obtain that

$$\partial_k \left(P_{ij} \frac{d_3}{1 + \nu_3} \right) = (\partial_k P_{ij}) \frac{d_3}{1 + \nu_3} + P_{ij} \frac{(\partial_k d_3)}{1 + \nu_3} - P_{ij} \frac{d_3 (\partial_k \nu_3)}{(1 + \nu_3)^2} =: T_1 + T_2 + T_3, \quad (85a)$$

$$\partial_k \left(\frac{\mu_j d_i}{1 + \nu_3} \right) = \frac{(\partial_k \mu_j) d_i}{1 + \nu_3} + \frac{\mu_j (\partial_k d_i)}{1 + \nu_3} - \frac{\mu_j d_i (\partial_k \nu_3)}{(1 + \nu_3)^2} =: T_4 + T_5 + T_6, \quad (85b)$$

$$\partial_k \left(\frac{\nu_i d_j}{1 + \nu_3} \right) = \frac{(\partial_k \nu_i) d_j}{1 + \nu_3} + \frac{\nu_i (\partial_k d_j)}{1 + \nu_3} - \frac{\nu_i d_j (\partial_k \nu_3)}{(1 + \nu_3)^2} =: T_7 + T_8 + T_9. \quad (85c)$$

With $1 + \mu_3 \geq \gamma > 0$ and $1 + \nu_3 \geq \gamma > 0$, Lemma A.1 (i)–(ii) yields that

$$\sum_{i \in \{1, 3, 4, 6, 7, 9\}} |T_i| \lesssim \gamma^{-2} \left(\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)} \right) \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}, \quad \text{and} \\ \sum_{i \in \{2, 5, 8\}} |T_i| \lesssim \gamma^{-1} \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}.$$

Together with (84), this proves (ii). For the proof of (iii), let $\ell, k \in \{1, 2, 3\}$. We differentiate the terms in (85) separately and exploit that the second derivative of piecewise

affine functions is zero. We start from (85a). Elementwise, the product rule yields that

$$\begin{aligned}\partial_\ell \left((\partial_k P_{ij}) \frac{d_3}{1 + \nu_3} \right) &= (\partial_\ell \partial_k P_{ij}) \frac{d_3}{1 + \nu_3} + (\partial_k P_{ij}) \frac{\partial_\ell d_3}{1 + \nu_3} - (\partial_k P_{ij}) \frac{d_3 (\partial_\ell \nu_3)}{(1 + \nu_3)^2}, \\ &=: \tilde{T}_1 + \tilde{T}_2 + \tilde{T}_3.\end{aligned}\quad (86a)$$

$$\partial_\ell \left(P_{ij} \frac{(\partial_k d_3)}{1 + \nu_3} \right) = (\partial_\ell P_{ij}) \frac{(\partial_k d_3)}{1 + \nu_3} - P_{ij} \frac{(\partial_k d_3)(\partial_\ell \nu_3)}{(1 + \nu_3)^2} =: \tilde{T}_4 + \tilde{T}_5, \quad (86b)$$

$$\begin{aligned}\partial_\ell \left(P_{ij} \frac{d_3 (\partial_k \nu_3)}{(1 + \nu_3)^2} \right) &= (\partial_\ell P_{ij}) \frac{d_3 (\partial_k \nu_3)}{(1 + \nu_3)^2} + P_{ij} \frac{(\partial_\ell d_3)(\partial_k \nu_3)}{(1 + \nu_3)^2} - 2P_{ij} \frac{d_3 (\partial_k \nu_3)(\partial_\ell \nu_3)}{(1 + \nu_3)^3} \\ &=: \tilde{T}_6 + \tilde{T}_7 + \tilde{T}_8.\end{aligned}\quad (86c)$$

Next, we get for the terms from (85b) elementwise that

$$\partial_\ell \left(\frac{(\partial_k \mu_j) d_i}{1 + \nu_3} \right) = \frac{(\partial_k \mu_j)(\partial_\ell d_i)}{1 + \nu_3} - \frac{(\partial_k \mu_j) d_i (\partial_\ell \nu_3)}{(1 + \nu_3)^2} =: \tilde{T}_9 + \tilde{T}_{10}, \quad (87a)$$

$$\partial_\ell \left(\frac{\mu_j (\partial_k d_i)}{1 + \nu_3} \right) = \frac{(\partial_\ell \mu_j)(\partial_k d_i)}{1 + \nu_3} - \frac{\mu_j (\partial_k d_i)(\partial_\ell \nu_3)}{(1 + \nu_3)^2} =: \tilde{T}_{11} + \tilde{T}_{12}, \quad (87b)$$

$$\begin{aligned}\partial_\ell \left(\frac{\mu_j d_i (\partial_k \nu_3)}{(1 + \nu_3)^2} \right) &= \frac{(\partial_\ell \mu_j) d_i (\partial_k \nu_3)}{(1 + \nu_3)^2} + \frac{\mu_j (\partial_\ell d_i)(\partial_k \nu_3)}{(1 + \nu_3)^2} - 2 \frac{\mu_j d_i (\partial_k \nu_3)(\partial_\ell \nu_3)}{(1 + \nu_3)^3} \\ &=: \tilde{T}_{13} + \tilde{T}_{14} + \tilde{T}_{15}.\end{aligned}\quad (87c)$$

Lemma A.1, $1 \leq 2/\gamma$ and $1 + \mu_3 \geq \gamma > 0$ as well as $1 + \nu_3 \geq \gamma > 0$ yield that

$$\begin{aligned}\sum_{i \in \{1, 3, 6, 8, 10, 12, 13, 15\}} |\tilde{T}_i| &\lesssim \gamma^{-3} (\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)})^2 \|\boldsymbol{\mu}_h - \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}, \quad \text{and} \\ \sum_{i \in \{2, 4, 5, 7, 9, 11, 14\}} |\tilde{T}_i| &\lesssim \gamma^{-2} (\|\nabla \boldsymbol{\mu}_h\|_{L^\infty(\Omega)} + \|\nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}) \|\nabla \boldsymbol{\mu}_h - \nabla \boldsymbol{\nu}_h\|_{L^\infty(\Omega)}.\end{aligned}$$

Note that the terms in (85c) are obtained if in (85b) we replace μ_i with ν_i and d_i with d_j . Hence, we can apply the same arguments as in (87). This proves (iii). \square

Lemma A.3 ([ABV14, Lemma 3.1]). *Let $\mathbf{B} \in \mathbb{R}^{2N \times 2N}$ be a positive definite matrix and $\mathbf{B}_0 \in \mathbb{R}^{2N \times 2N}$ be a symmetric positive definite matrix, which satisfy for $c_1, c_2 > 0$ that*

$$\begin{aligned}\mathbf{x} \cdot \mathbf{B} \mathbf{x} &\geq c_1 \mathbf{x} \cdot \mathbf{B}_0 \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N} \quad \text{and} \\ \mathbf{x} \cdot \mathbf{B} \mathbf{y} &\leq c_2 (\mathbf{x} \cdot \mathbf{B}_0 \mathbf{x})^{1/2} (\mathbf{y} \cdot \mathbf{B}_0 \mathbf{y})^{1/2} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}.\end{aligned}$$

Then, it holds that

$$\begin{aligned}\mathbf{x} \cdot \mathbf{B}^{-1} \mathbf{x} &\geq \frac{c_1}{c_2^2} \mathbf{x} \cdot \mathbf{B}_0^{-1} \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^{2N} \quad \text{and} \\ \mathbf{x} \cdot \mathbf{B}^{-1} \mathbf{y} &\leq c_1^{-1} (\mathbf{x} \cdot \mathbf{B}_0^{-1} \mathbf{x})^{1/2} (\mathbf{y} \cdot \mathbf{B}_0^{-1} \mathbf{y})^{1/2} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^{2N}.\end{aligned}\quad \square$$

REFERENCES

- [ABV14] B. Ayuso de Dios, A. T. Barker, and P. S. Vassilevski. A combined preconditioning strategy for nonsymmetric systems. *SIAM J. Sci. Comput.*, 36(6):A2533–A2556, 2014.
- [AHP⁺14] C. Abert, G. Hrkac, M. Page, D. Praetorius, M. Ruggeri, and D. Suess. Spin-polarized transport in ferromagnetic multilayers: an unconditionally convergent FEM integrator. *Comput. Math. Appl.*, 68(6):639–654, 2014.
- [AJ06] F. Alouges and P. Jaisson. Convergence of a finite element discretization for the Landau-Lifshitz equations in micromagnetism. *Math. Models Methods Appl. Sci.*, 16(2):299–316, 2006.
- [AKST14] F. Alouges, E. Kritsikis, J. Steiner, and J.-C. Toussaint. A convergent and precise finite element scheme for Landau-Lifshitz-Gilbert equation. *Numer. Math.*, 128(3):407–430, 2014.
- [AKT12] F. Alouges, E. Kritsikis, and J.-C. Toussaint. A convergent finite element approximation for Landau-Lifshitz-Gilbert equation. *Physica B*, 407:1345–1349, 2012.
- [Alo08] F. Alouges. A new finite element scheme for Landau-Lifshitz equations. *Discrete Contin. Dyn. Syst. Ser. S*, 1(2):187–196, 2008.
- [ARB⁺15] C. Abert, M. Ruggeri, F. Bruckner, C. Vogler, G. Hrkac, D. Praetorius, and D. Suess. A three-dimensional spin-diffusion model for micromagnetics. *Scientific Reports*, (5):14855, 2015.
- [AS92] F. Alouges and A. Soyeur. On global weak solutions for Landau-Lifshitz equations: existence and nonuniqueness. *Nonlinear Anal.*, 18(11):1071–1084, 1992.
- [Bar15] S. Bartels. *Numerical methods for nonlinear partial differential equations*, volume 47 of *Springer Series in Computational Mathematics*. Springer, Cham, 2015.
- [BKP08] S. Bartels, J. Ko, and A. Prohl. Numerical analysis of an explicit approximation scheme for the Landau-Lifshitz-Gilbert equation. *Math. Comp.*, 77(262):773–788, 2008.
- [BPP15] L. Bañas, M. Page, and D. Praetorius. A convergent linear finite element scheme for the Maxwell-Landau-Lifshitz-Gilbert equations. *Electron. Trans. Numer. Anal.*, 44:250–270, 2015.
- [BPPR14] L. Bañas, M. Page, D. Praetorius, and J. Rochat. A decoupled and unconditionally convergent linear FEM integrator for the Landau-Lifshitz-Gilbert equation with magnetostriction. *IMA J. Numer. Anal.*, 34(4):1361–1385, 2014.
- [BSF⁺14] F. Bruckner, D. Suess, M. Feischl, T. Führer, P. Goldenits, M. Page, D. Praetorius, and M. Ruggeri. Multiscale modeling in micromagnetics: existence of solutions and numerical integration. *Math. Models Methods Appl. Sci.*, 24(13):2627–2662, 2014.
- [DPP⁺17] G. Di Fratta, C.-M. Pfeiler, D. Praetorius, M. Ruggeri, and B. Stiftner. Linear second order IMEX-type integrator for the (eddy current) Landau-Lifshitz-Gilbert equation, 2017. Extended preprint available at arXiv:1711.10715.
- [DS14] E. Dumas and F. Sueur. On the weak solutions to the Maxwell-Landau-Lifshitz equations and to the Hall-Magneto-Hydrodynamic equations. *Commun. Math. Phys.*, 330:1179–1225, 2014.
- [FT17] M. Feischl and T. Tran. The eddy current-LLG equations: FEM-BEM coupling and a priori error estimates. *SIAM J. Numer. Anal.*, 55(4):1786–1819, 2017.
- [Gil55] T. L. Gilbert. A Lagrangian formulation of the gyromagnetic equation of the magnetization fields. *Phys. Rev.*, 100:1243, 1955. Abstract only.
- [gmr] Iterative template routine – GMRES. www.netlib.org/templates/cpp/gmres.h. Accessed: 2018-08-21.
- [GO95] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70(2):163–180, 1995.
- [HPP⁺18] G. Hrkac, C.-M. Pfeiler, D. Praetorius, M. Ruggeri, A. Segati, and B. Stiftner. Convergent tangent plane integrators for the simulation of chiral magnetic skyrmion dynamics, 2018. Accepted for publication in *Adv. Comput. Math.*, preprint available at arXiv:1712.03795v2.
- [HS98] A. Hubert and R. Schäfer. *Magnetic domains; the analysis of magnetic microstructures*. Springer, Berlin, 1998. Corrected Printing 2000.

- [LL08] L. Landau and E. Lifschitz. On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. *Ukr. J. Phys.*, (53), 2008. Special Issue, Reprinted from *Phys. Zeitsch. der Sow.* 8, pp. 153–169 (1935).
- [LPPT15] K.-N. Le, M. Page, D. Praetorius, and T. Tran. On a decoupled linear FEM integrator for eddy-current-LLG. *Appl. Anal.*, 94(5):1051–1067, 2015.
- [LT13] K.-N. Le and T. Tran. A convergent finite element approximation for the quasi-static Maxwell-Landau-Lifshitz-Gilbert equations. *Comput. Math. Appl.*, 66(8):1389–1402, 2013.
- [mum] μ MAG – Micromagnetic Modeling Activity Group. National Institute for Standards and Technology (NIST). <http://www.ctcms.nist.gov/~rdm/mumag.org.html>. Accessed: 2016-10-25.
- [Nep91] S. V. Nepomnyaschikh. Mesh theorems on traces, normalizations of function traces and their inversion. *Soviet J. Numer. Anal. Math. Modelling*, 6(3):223–242, 1991.
- [ngs] NGSolve finite element library. <https://ngsolve.org/>. Accessed: 2018-04-19.
- [NS96] S. G. Nash and A. Sofer. Preconditioning reduced matrices. *SIAM J. Matrix Anal. Appl.*, 17(1):47–68, 1996.
- [PRS18] D. Praetorius, M. Ruggeri, and B. Stiftner. Convergence of an implicit–explicit midpoint scheme for computational micromagnetics. *Comput. Math. Appl.*, 75(5):1719–1738, 2018.
- [Rie] A. Rieder. Coupling NGSolve to BEM++. <https://github.com/ariieder/ngbem>. Accessed: 2018-02-27.
- [Rug16] M. Ruggeri. *Coupling and numerical integration of the Landau–Lifshitz–Gilbert equation*. PhD thesis, TU Wien, Institute for Analysis and Scientific Computing, 2016.
- [Saa03] Yousef Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.
- [ŠBA+15] W. Šmigaj, T. Betcke, S. Arridge, J. Phillips, and M. Schweiger. Solving boundary integral problems with BEM++. *ACM Trans. Math. Software*, 41(2):Art. 6, 40, 2015.
- [SS86] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7(3):856–869, 1986.
- [Sta97] G. Starke. Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems. *Numer. Math.*, 78(1):103–117, 1997.
- [TNMS05] A. Thiaville, Y. Nakatani, J. Miltat, and Y. Suzuki. Micromagnetic understanding of current-driven domain wall motion in patterned nanowires. *EPL (Europhysics Letters)*, 69(6):990, 2005.
- [ZL04] S. Zhang and Z. Li. Roles of nonequilibrium conduction electrons on the magnetization dynamics of ferromagnets. *Phys. Rev. Lett.*, 93(12):127204, 2004.

FACULTY OF MATHEMATICS, UNIVERSITY OF DUISBURG-ESSEN, THEA-LEYMANN-STRASSE 9, 45127 ESSEN, GERMANY

E-mail address: johannes.kraus@uni-due.de

INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING, TU WIEN, WIEDNER HAUPTSTRASSE 8-10, 1040 VIENNA, AUSTRIA

E-mail address: carl-martin.pfeiler@asc.tuwien.ac.at

E-mail address: dirk.praetorius@asc.tuwien.ac.at

E-mail address: bernhard.stiftner@asc.tuwien.ac.at

FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, 1090 VIENNA, AUSTRIA

E-mail address: michele.ruggeri@univie.ac.at