APPLICATIONS OF $\mathcal{H}$-MATRIX TECHNIQUES IN MICROMAGNETICS

NIKOLA POPOVIĆ AND DIRK PRAETORIUS

ABSTRACT. The variational model by LANDAU and LIFSHITZ is frequently used in the simulation of stationary micromagnetic phenomena. We consider the limit case of large and soft magnetic bodies, treating the associated Maxwell equation exactly via an integral operator $\mathcal{P}$. In numerical simulations of the resulting minimization problem, difficulties arise due to the imposed side-constraint and the unboundedness of the domain. We introduce a possible discretization by a penalization strategy. Here, the computation of $\mathcal{P}$ is numerically the most challenging issue, as it leads to densely populated matrices. We show how an efficient treatment of both $\mathcal{P}$ and the corresponding bilinear form can be achieved by application of $\mathcal{H}$-matrix techniques.

1. INTRODUCTION

The simulation of stationary micromagnetic phenomena occurring in static or quasi-static processes is frequently based on a variational model named after LANDAU and LIFSHITZ. Therein, one minimizes the energy functional

$$E_\alpha(\mathbf{m}) := \int_\Omega \phi(\mathbf{m})\,dx - \int_\Omega \mathbf{f}\cdot\mathbf{m}\,dx + \frac{1}{2}\int_{\mathbb{R}^d} |\nabla u|^2\,dx + \alpha\int_\Omega |\nabla\mathbf{m}|^2\,dx \tag{1}$$

over some set of admissible vector-valued magnetizations $\mathbf{m}:\Omega\to\mathbb{R}^d$ on a bounded Lipschitz domain $\Omega\subset\mathbb{R}^d$ corresponding to the magnet, with $\mathbf{m}(x) := 0$ for $x\in\mathbb{R}^d\backslash\overline{\Omega}$ and $d = 2, 3$. Here, $\phi\in\mathcal{C}^\infty(\mathbb{R}^d;\mathbb{R}^+)$ is the *anisotropy density* (depending on properties of the material on a crystalline level), $\mathbf{f}\in L^2(\Omega;\mathbb{R}^d)$ denotes an applied *exterior magnetic field*, $0\le\alpha\ll 1$ is the *exchange parameter*, and u is the *magnetic potential* related to $\mathbf{m}$ by Maxwell's equation

$$\operatorname{div}(-\nabla u + \mathbf{m}) = 0 \quad \text{in } \mathcal{D}'(\mathbb{R}^d). \tag{2}$$

The model is completed by adding the non-convex constraint

$$|\mathbf{m}(x)| = 1 \quad \text{for a.e. } x\in\Omega. \tag{3}$$

For large and soft magnets, the parameter α in (1) vanishes. In general, the model then lacks classical solutions, see [JK90], and hence has to be relaxed either by considering measure valued solutions [Ped97] or by convexification [DeS93, Tar95]. In fact, for a certain limit configuration of soft-large bodies, $E_0(\mathbf{m})$, i.e., (1) with $\alpha\to 0$ can be justified to be the correct model, see [DeS93]. The corresponding

1991 *Mathematics Subject Classification.* 65K10, 65F30, 65F50, 65N30, 65N38.

Key words and phrases. Micromagnetics; Landau-Lifshitz model; Magnetic potential; Finite element method; Hierarchical matrices; Interpolation.

convexified problem E_0^{**} is given by

$$E_0^{**}(\mathbf{m}) := \int_\Omega \phi^{**}(\mathbf{m})\,dx - \int_\Omega \mathbf{f}\cdot\mathbf{m}\,dx + \frac{1}{2}\int_{\mathbb{R}^d} |\nabla u|^2\,dx \tag{4}$$

subject to (2) and

$$|\mathbf{m}(x)| \le 1 \quad \text{for a.e. } x \in \Omega. \tag{5}$$

Here, ϕ^{**} is the convexified density defined by

$$\phi^{**}(x) = \sup\big\{\varphi(x) \,\big|\, \varphi : \mathbb{R}^d \to \mathbb{R} \text{ convex and } \varphi|_{\mathbb{S}} \le \phi\big\} \quad \text{for } |x| \le 1, \tag{6}$$

where $\mathbb{S} = \big\{x \in \mathbb{R}^d \,\big|\, |x| = 1\big\}$ denotes the unit sphere. Then, the relaxed problem reads:

$$\text{Minimize } E_0^{**} \text{ over } \mathcal{A} := \big\{\mathbf{m} \in L^\infty(\Omega;\mathbb{R}^d) \,\big|\, \|\mathbf{m}\|_{L^\infty(\Omega;\mathbb{R}^d)} \le 1\big\}. \tag{7}$$

In contrast to the ill-posed problem E_0, the convexification is well-posed [DeS93, Ped97, CP04b]. In fact, this convexified model provides the mathematical foundation of the so-called *phase theory* in micromagnetics, cf. [HS98].

***Remark* 1.** For *uniaxial* materials such as cobalt, the anisotropy energy is given by $\phi(x) = 1/2\big(1-(x\cdot\mathbf{e})^2\big)$, with $|x| = 1$ and $\mathbf{e} \in \mathbb{R}^d$ a given fixed unit vector called the *easy axis*. A direct calculation shows $\phi^{**}(x) = 1/2\sum_{j=2}^d (x\cdot\mathbf{z}_j)^2$ for $|x| \le 1$ then, where $\{\mathbf{e}, \mathbf{z}_2, \dots, \mathbf{z}_d\}$ is an orthonormal basis of $\mathbb{R}^d$.

The numerical treatment of the minimization problem related to E_0^{**} was initiated by [CP01] for $d = 2$, where the authors treat a simplified model obtained by replacing $\mathbb{R}^d$ in (2) by a bounded Lipschitz domain $\widehat{\Omega}$ containing Ω, and solve for a potential $u \in H_0^1(\widehat{\Omega})$. Here, as in [CP04b, CP04a, LM92, Ma91, Pra04], (2) is treated *exactly* via an integral representation, i.e., $u = \mathcal{L}\mathbf{m}$, where $\mathcal{L}$ is a linear *convolution operator*. We then set $\mathcal{P}\mathbf{m} := \nabla(\mathcal{L}\mathbf{m})$, see Theorem 2.1 below, and reformulate the stray field energy contribution in (4) in terms of $\mathcal{P}$. The advantage is that in the resulting model, only one discretization for $\mathbf{m}$ is required, e.g., by piecewise constant functions $\mathbf{m}_h$.
From a numerical point of view, the computation of $\mathcal{P}\mathbf{m}$ for a given magnetization is the most challenging issue, since it will lead to densely populated matrices. The aim of the present work is to show how an efficient numerical treatment of both $\mathcal{P}$ and the induced bilinear form $a(\cdot,\cdot)$ can be achieved by application of $\mathcal{H}$-matrix techniques.

***Remark* 2.** The treatment of the convexified model (7) requires the explicit knowledge of the convexified anisotropy density ϕ^{**}, which is, however, in general unknown even for simple ϕ, cf. [DeS93]. Both the Young measure relaxation proposed in [Ped97] and the corresponding discretization [KP01] avoid the computation of ϕ^{**}. Note that as far as the computation of the magnetic potential (2) is concerned, our ideas apply in that setting, as well. Further analysis on stabilized discrete models, as well as a comparison of the various approaches, can be found in the articles [CP01] and [KP01], as well as in the survey monograph by Prohl [Pro01].

The remainder of this paper is organized as follows: in Section 2, we give a few preliminaries and present a possible discretization of (7); Section 3 contains some interpolation results required for the following analysis; Section 4 motivates the

concept of hierarchical ($\mathcal{H}$- resp. $\mathcal{H}^2$-) matrices; in Section 5, we give two different approaches for a Galerkin discretization of the potential equation (2) via $\mathcal{H}$-matrix techniques; Section 6 finally summarizes the results of our numerical experiments.

2. Preliminaries and Discretization

This section is devoted to the reformulation of (7) in terms of the associated Euler-Lagrange equations and introduces a possible discretization by a penalization strategy.

2.1. Preliminaries. The following Theorem 2.1 gathers some of the properties of the operator $\mathcal{P}$ required in the following. Proofs can be found in [Pra04], although we expect the result to be known to the experts.

Theorem 2.1 ([LM92, Ma91, Pra04])**.** *Given any $\mathbf{m} \in L^\infty(\Omega;\mathbb{R}^d)$, there exists an (up to an additive constant) unique magnetic potential $u = \mathcal{L}\mathbf{m} \in H^1_{\ell oc}(\mathbb{R}^d)$ such that*

$$\nabla u \in L^2(\mathbb{R}^d;\mathbb{R}^d) \quad \text{and} \quad \operatorname{div}(-\nabla u + \mathbf{m}) = 0 \quad \text{in } \mathcal{D}'(\mathbb{R}^d). \tag{8}$$

The (extended) operator $\mathcal{P} : L^2(\mathbb{R}^d;\mathbb{R}^d) \to L^2(\mathbb{R}^d;\mathbb{R}^d)$, $\mathbf{m} \mapsto \nabla(\mathcal{L}\mathbf{m})$ is an L^2 orthogonal projection. The potential $\mathcal{L}\mathbf{m}$ can be represented as a convolution operator

$$\mathcal{L}\mathbf{m} = \sum_{j=1}^{d} \frac{\partial G}{\partial x_j} * \mathbf{m}_j, \tag{9}$$

where $\mathbf{m} = (\mathbf{m}_1, \dots, \mathbf{m}_d)$ is trivially extended [by zero] from Ω to $\mathbb{R}^d$ (so that the convolution is formally well-defined). Here $G : \mathbb{R}^d\backslash\{0\} \to \mathbb{R}$ is the Newtonian kernel

$$G(x) := \begin{cases} \dfrac{1}{\gamma_2} \log|x|, & d = 2, \\ \dfrac{1}{(2-d)\gamma_d} |x|^{2-d}, & d > 2 \end{cases} \tag{10}$$

for $x \neq 0$, where the constant $\gamma_d := |\mathbb{S}| > 0$ denotes the surface measure of the unit sphere (e.g., $\gamma_2 = 2\pi$, $\gamma_3 = 4\pi$). □

Since the energy functional E_0^{**} from (4) is convex and (Gâteaux) differentiable, the minima are equivalently characterized by the corresponding Euler-Lagrange equations [DeS93]. Thus, problem (RP) reads: find $(\lambda, \mathbf{m}) \in L^2(\Omega) \times L^2(\Omega;\mathbb{R}^d)$ such that

$$\mathcal{P}\mathbf{m} + D\phi^{**}(\mathbf{m}) + \lambda\mathbf{m} = \mathbf{f} \quad \text{a.e. in } \Omega, \tag{11}$$

$$\lambda \geq 0,\ |\mathbf{m}| \leq 1,\ \lambda(1 - |\mathbf{m}|) = 0 \quad \text{a.e. in } \Omega. \tag{12}$$

Existence results for (RP) can be found in [DeS93, Ped97]; in particular, in the uniaxial case the solution to (RP) is unique.

2.2. The Discretized Problem. Let $\mathcal{T} = \{T_1, \dots, T_N\}$ be a finite family of pairwise disjoint non-empty open sets T_j which satisfy $\overline{\Omega} = \bigcup_{j=1}^N \overline{T_j}$. The space of all $\mathcal{T}$-piecewise constant functions is denoted by $\mathcal{P}_0(\mathcal{T})$; $h \in \mathcal{P}_0(\mathcal{T})$ is the mesh-size

function, $h|_T := h_T := \operatorname{diam}(T)$. For $f \in L^2(\Omega)$, let $f_{\mathcal{T}} \in \mathcal{P}_0(\mathcal{T})$ be the $\mathcal{T}$-piecewise integral mean given by

$$f_{\mathcal{T}}|_T := \frac{1}{|T|}\int_T f\,dx \qquad \text{for all } T \in \mathcal{T}.$$

The discrete problem $(RP_{\varepsilon,h})$ now reads as follows: given a penalization parameter $\varepsilon \in \mathcal{P}_0(\mathcal{T})$ with $\varepsilon > 0$, find $\mathbf{m}_h \in \mathcal{P}_0(\mathcal{T})^d$ such that

(13)
$$\langle \mathcal{P}\mathbf{m}_h + D\phi^{**}(\mathbf{m}_h) + \lambda_h \mathbf{m}_h \,;\, \widetilde{\mathbf{m}}_h\rangle_{L^2(\Omega)} = \langle \mathbf{f} \,;\, \widetilde{\mathbf{m}}_h\rangle_{L^2(\Omega)} \quad \text{for all } \widetilde{\mathbf{m}}_h \in \mathcal{P}_0(\mathcal{T})^d,$$

where $\lambda_h \in \mathcal{P}_0(\mathcal{T})$ is defined by

$$\lambda_h = \varepsilon^{-1}\frac{(|\mathbf{m}_h| - 1)_+}{|\mathbf{m}_h|} \tag{14}$$

with $(\cdot)_+ := \max\{\cdot, 0\}$.

***Remark* 3.** For $\mathbf{m}_h \in \mathcal{P}_0(\mathcal{T})^d$, the potential $\mathcal{P}\mathbf{m}_h$ can be computed exactly, as the associated bilinear form

$$a(\mathbf{m}_h, \widetilde{\mathbf{m}}_h) := \langle \mathcal{P}\mathbf{m}_h \,;\, \widetilde{\mathbf{m}}_h\rangle_{L^2(\Omega)} \quad \text{for all } \mathbf{m}_h, \widetilde{\mathbf{m}}_h \in \mathcal{P}_0(\mathcal{T})^d \tag{15}$$

can be evaluated by a closed form formula, cf. Theorem 5.1. The evaluation of $\mathcal{P}\mathbf{m}_h$ is, however, computationally demanding, as it typically leads to densely populated stiffness matrices.

As for the continuous problem, we have existence of discrete solutions and uniqueness in the uniaxial model case, cf. [CP04b]. Moreover, in that case the a priori error analysis for $(RP_{\varepsilon,h})$ from [CP04b] suggests the choice $\varepsilon = h$ for the penalization parameter, as there holds

$$\begin{aligned}\|\mathcal{P}\mathbf{m} - \mathcal{P}\mathbf{m}_h\|_{L^2(\mathbb{R}^d)} &+ \|D\phi^{**}(\mathbf{m}) - D\phi^{**}(\mathbf{m}_h)\|_{L^2(\Omega)} \\ &+ \|\lambda\mathbf{m} - \lambda_h\mathbf{m}_h\|_{L^2(\Omega)} + \|\varepsilon\lambda_h\mathbf{m}_h\|_{L^2(\Omega)} \\ &\le C\Big(\|\mathbf{m} - \mathbf{m}_{\mathcal{T}}\|_{L^2(\Omega)} + \|\lambda\mathbf{m} - (\lambda\mathbf{m})_{\mathcal{T}}\|_{L^2(\Omega)} + \|\varepsilon\lambda\mathbf{m}\|_{L^2(\Omega)}\Big)\end{aligned}$$

with a generic constant $C \ge 0$. For $(\lambda, \mathbf{m})$ sufficiently smooth (e.g., $\mathbf{m} \in H^1(\Omega;\mathbb{R}^d)$, $\lambda\mathbf{m} \in H^1(\Omega;\mathbb{R}^d)$), the above right-hand side turns out to be of order $\mathcal{O}(\varepsilon + h)$.

The stiffness matrix $\mathbf{A}$ induced by the bilinear form $a(\cdot,\cdot)$ from (15) will in the following be approximated by an appropriate $\mathcal{H}^2$-matrix $\widetilde{\mathbf{A}}$. Given $(RP_{\varepsilon,h})$, one then obtains an approximate discrete model $(\widetilde{RP}_{\varepsilon,h})$ after replacing $\mathbf{A}$ by $\widetilde{\mathbf{A}}$ and defining the approximate bilinear form $\widetilde{a}(\cdot,\cdot)$ accordingly.

3. Multidimensional Interpolation of Integral Kernels

The following section contains some results on the multidimensional interpolation of integral kernels. We restrict ourselves to one particular class of kernel functions here, known as asymptotically smooth kernels.

3.1. **Asymptotically Smooth Kernels.** A kernel function

$$\kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \quad (x,y) \mapsto \kappa(x,y) \tag{16}$$

is said to be *asymptotically smooth* if there exist constants C_{asm} and c_{asm} such that

$$|\partial_x^\alpha \partial_y^\beta \kappa(x,y)| \leq C_{\mathrm{asm}}(c_{\mathrm{asm}}|x-y|)^{-|\alpha|-|\beta|-s}(\alpha+\beta)! \tag{17}$$

for all multi-indices $\alpha, \beta \in \mathbb{N}_0^d$ with $|\alpha| + |\beta| \geq 1$ and some singularity order $s \in \mathbb{R}$, where $x, y \in \mathbb{R}^d$ with $x \neq y$.

Example 3.1. *For the Newtonian kernel G defined in* (10)*, $\kappa(x,y) := G(x-y)$ is asymptotically smooth for any $d \geq 2$, with $C_{\mathrm{asm}} = \gamma_2^{-1}$ and $C_{\mathrm{asm}} = \big((2-d)\gamma_d\big)^{-1}$ for $d = 2$ and $d \geq 3$, respectively, and $c_{\mathrm{asm}} = 1$, see* [Gra01].

***Remark* 4.** Note that the derivatives of an asymptotically smooth kernel κ also are asymptotically smooth: given $\tilde{\kappa} := \partial_x^{\tilde{\alpha}} \partial_y^{\tilde{\beta}} \kappa$, (17) holds with $\tilde{s} = s + |\tilde{\alpha}| + |\tilde{\beta}|$, $\tilde{c}_{\mathrm{asm}} = c_{\mathrm{asm}} + \varepsilon$ for $\varepsilon > 0$ arbitrary, and some $\widetilde{C}_{\mathrm{asm}}$ depending on C_{asm} and ε. This is a consequence of

$$\big((\alpha+\beta)+(\tilde{\alpha}+\tilde{\beta})\big)! \leq \overline{C}_{\mathrm{asm}}\, \overline{c}_{\mathrm{asm}}^{|\alpha|+|\beta|}(\alpha+\beta)!$$

with an $(|\tilde{\alpha}| + |\tilde{\beta}|)$-dependent constant $\overline{c}_{\mathrm{asm}}$. For every choice of $\overline{c}_{\mathrm{asm}} > 1$, there is a $\overline{C}_{\mathrm{asm}} > 0$ (depending on $\overline{c}_{\mathrm{asm}}$) such that the above inequality holds. One then sets $\widetilde{C}_{\mathrm{asm}} := C_{\mathrm{asm}}\overline{C}_{\mathrm{asm}}$ and $\widetilde{c}_{\mathrm{asm}} := c_{\mathrm{asm}}\overline{c}_{\mathrm{asm}}$, respectively.

3.2. **Interpolation Operators in One Dimension.** For $m \in \mathbb{N}_0$, let the space of m-th order polynomials in one spatial variable be denoted by $\mathcal{P}_m$, and consider the *interpolation operator*

$$\mathcal{I}_m : \mathcal{C}[-1,1] \to \mathcal{P}_m, \quad u \mapsto \sum_{j=0}^{m} u(t_j)\mathcal{L}_j \quad \text{with } \mathcal{L}_j(t) = \prod_{\substack{k=0 \\ k\neq j}}^{m} \frac{t-t_k}{t_j-t_k} \tag{18}$$

acting on the so-called *reference element* $[-1,1]$. Here $\big(\mathcal{L}_j(t)\big)_{j=0}^m$ are the Lagrange polynomials corresponding to the interpolation points $(t_j)_{j=0}^m$. Note that $\mathcal{I}_m$ is a projection, i.e., linear with $\mathcal{I}_m^2 = \mathcal{I}_m$.

For $m \in \mathbb{N}_0$, the *Lebesgue constant* $\Lambda_m \in \mathbb{R}$ is defined as the operator norm of $\mathcal{I}_m$,

$$\Lambda_m := \sup_{\substack{u \in \mathcal{C}[-1,1] \\ u \neq 0}} \frac{\|\mathcal{I}_m u\|_{\infty,[-1,1]}}{\|u\|_{\infty,[-1,1]}}. \tag{19}$$

Clearly, we have $\Lambda_m \geq 1$. Moreover, we assume that there are constants $\lambda, C_\lambda \in \mathbb{R}^+$ such that

$$\Lambda_m \leq C_\lambda (m+1)^\lambda. \tag{20}$$

For *Chebyshev interpolation*, where $t_j = \cos\big((2j+1)\pi/(2(m+1))\big)$, this estimate holds with $\lambda = 1 = C_\lambda$, cf. [Riv84].

For an arbitrary compact interval $I := [a,b] \subset \mathbb{R}$, we define the affine transformation

$$\Phi_I : [-1,1] \to I, \quad t \mapsto \frac{1}{2}\big((a+b) + t(b-a)\big).$$

The transformed interpolation operator $\mathcal{I}_m^I$ is then given by

$$\mathcal{I}_m^I : \mathcal{C}[a,b] \to \mathcal{P}_m, \quad u \mapsto \big(\mathcal{I}_m(u \circ \Phi_I)\big) \circ \Phi_I^{-1}.$$

Obviously, the projection property as well as (20) now carry over from $\mathcal{I}_m$ to $\mathcal{I}_m^I$.

3.3. Tensor Interpolation Operators. For a family of closed intervals $I_j := [a_j, b_j] \subset \mathbb{R}$, $j \in \{1, \dots, 2d\}$, define the axially parallel box $B \subset \mathbb{R}^{2d}$ by $B := \prod_{j=1}^{2d} I_j$. Given a family $\big(\mathcal{I}_m^{I_j}\big)_{j=0}^{2d}$ of interpolation operators on $(I_j)_{j=0}^{2d}$, the m-th order *tensor product interpolation operator* $\mathcal{I}_m^B$ on B is then defined as

$$\mathcal{I}_m^B := \mathcal{I}_m^{I_1} \otimes \cdots \otimes \mathcal{I}_m^{I_{2d}}.$$

In analogy to $\mathcal{I}_m^I$, $\mathcal{I}_m^B$ is a projection from $\mathcal{C}(B)$ to

$$\mathcal{Q}_m := \operatorname{span}\big\{p_1 \otimes \cdots \otimes p_{2d} \,\big|\, p_j \in \mathcal{P}_m, j \in \{1, \dots, 2d\}\big\}. \tag{21}$$

We require the following result on the interpolation error of $\mathcal{I}_m^B$ adapted from [BG04, BLM04]:

Theorem 3.2. *Let $u \in \mathcal{C}^\infty(B)$ such that there are constants $C_u, \gamma_u \in \mathbb{R}^+$ satisfying*

$$\|\partial_j^n u\|_{\infty,B} \le C_u \gamma_u^n n! \tag{22}$$

for all $j \in \{1, \dots, 2d\}$ and $n \in \mathbb{N}_0$. Then, we have

(23)

$$\|u - \mathcal{I}_m^B u\|_{\infty,B} \le 16 e d C_u \Lambda_m^{2d} \big(1 + \gamma_u \operatorname{diam}(B)\big)(m+1) \left(1 + \frac{2}{\gamma_u \operatorname{diam}(B)}\right)^{-(m+1)}.$$

Proof. The proof is as in [BG04, Theorem 3.2], with the d there replaced by $2d$. □

3.4. Local Error Analysis for Asymptotically Smooth Kernels. We now apply interpolation to obtain an approximate degenerate kernel $\widetilde{\kappa} := \mathcal{I}_m^B \kappa$ instead of the given asymptotically smooth integral kernel κ. Let $B_\sigma, B_\tau \subset \mathbb{R}^d$ be compact axially parallel boxes with positive Euclidean distance $\operatorname{dist}(B_\sigma, B_\tau) > 0$:

Lemma 3.3. *An asymptotically smooth kernel $\kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies* (22) *on $B := B_\sigma \times B_\tau$, with constants*

(24)

$$C_\kappa = \max\left\{ \|\kappa\|_{L^\infty(B_\sigma \times B_\tau)}, \frac{C_{\mathrm{asm}}}{\big(c_{\mathrm{asm}} \operatorname{dist}(B_\sigma, B_\tau)\big)^s} \right\} \quad \text{and} \quad \gamma_\kappa = \frac{1}{c_{\mathrm{asm}} \operatorname{dist}(B_\sigma, B_\tau)}.$$

Provided $\operatorname{diam}(B_\sigma \times B_\tau) \le \eta \operatorname{dist}(B_\sigma, B_\tau)$ *with $\eta > 0$, there holds in particular*

$$\|\kappa - \mathcal{I}_m^{(\sigma,\tau)} \kappa\|_{L^\infty(B_\sigma \times B_\tau)} \le c_1 c_2 C_\kappa \left(1 + \frac{2 c_{\mathrm{asm}}}{\eta}\right)^{-(m+1)} \tag{25}$$

with $\mathcal{I}_m^{(\sigma,\tau)} := \mathcal{I}_m^B$, a numerical constant $c_1 = 16 e d (1 + \eta / c_{\mathrm{asm}})$, and a constant $c_2 = \Lambda_m^{2d}(m+1)$ with only polynomial increase in m.

***Remark* 5.** Note that the constant $C_\kappa > 0$ behaves like $\operatorname{dist}(B_\sigma, B_\tau)^{-s}$ for the kernels we are interested in, such as $\kappa(x,y) = \log|x-y|$ resp. $\kappa(x,y) = |x-y|^{-s}$.

Proof of Lemma 3.3. Direct computation shows that (22) is valid, and Theorem 3.2 yields

$$\|\kappa - \mathcal{I}_m^B \kappa\|_{L^\infty(B)} \le 16 e d \big(1 + \gamma_\kappa \operatorname{diam}(B)\big) \Lambda_m^{2d} (m+1) C_\kappa \left(1 + \frac{2}{\gamma_\kappa \operatorname{diam}(B)}\right)^{-(m+1)}.$$

Combining this with $\gamma_\kappa \operatorname{diam}(B) \le \eta / c_{\mathrm{asm}}$, we obtain (25). □

The proof of Theorem 3.2 is only based on the stability constant Λ_m defined in (19). The advantage is that the resulting estimate can be applied to a fairly wide range of interpolation operators. If one restricts oneself to tensor Chebyshev interpolation – as we will do in the numerical experiments – one can do better by using the following error estimate for tensor Chebyshev polynomials adapted from [BH02]:

Lemma 3.4. *Provided* $\operatorname{diam}(B_\sigma \times B_\tau) \leq \eta \operatorname{dist}(B_\sigma, B_\tau)$ *with* $\eta > 0$*, an asymptotically smooth kernel* κ *on* $B := B_\sigma \times B_\tau \subset \mathbb{R}^{2d}$ *satisfies*

$$\|\kappa - \mathcal{I}_m^{(\sigma,\tau)}\kappa\|_{\infty,B} \leq dC_{\mathrm{asm}}c_{\mathrm{asm}}^{-s}\,\Lambda_m^{2d-1}\operatorname{dist}(B_\sigma, B_\tau)^{-s}\,4^{-m}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}. \tag{26}$$

Proof. The proof is along the lines of [BH02]. □

***Remark* 6.** Lemma 3.3 ensures (asymptotically) exponential convergence with respect to m irrespective of the choice of $\eta > 0$. Nevertheless, the constant c_1 obtained in Lemma 3.3 is too pessimistic, in contrast to the reasonably good approximation results observed for small m, as well, cf. Section 6. From Lemma 3.4 we obtain exponential convergence provided at least $4c_{\mathrm{asm}} > \eta$, with the highly improved constants $c_1 = dC_{\mathrm{asm}}c_{\mathrm{asm}}^{-s} \ll 16ed(1 + \eta/c_{\mathrm{asm}})$ and $c_2 = \Lambda_m^{2d-1} \ll \Lambda_m^{2d}(m+1)$.

In the following, we require additional error estimates for tensor Chebyshev polynomials, in particular for the norms of the first and second derivatives of the respective interpolation errors. For the proofs, we make use of the following well-known one-dimensional error estimate for Chebyshev interpolation from [BH02],

$$\|u - \mathcal{I}_m^{I_j}u\|_{\infty,I_j} \leq \frac{4^{-m}}{2(m+1)!}\,|I_j|^{m+1}\,\|u^{(m+1)}\|_{\infty,I_j} \quad \text{for } u \in \mathcal{C}^{m+1}(I_j), \tag{27}$$

as well as of a result on the first derivative of $u - \mathcal{I}_m u$ in one dimension taken from the proof of [BS01, Theorem 3.3.1],

$$\|(u - \mathcal{I}_m u)'\|_{\infty,[-1,1]} \leq \left(\frac{1}{(r-1)!} + \frac{1}{r!}C(m)\right)\|u^{(r)}\|_{\infty,[-1,1]} \quad \text{for } u \in \mathcal{C}^r[-1,1], \tag{28}$$

with $1 \leq r \leq m+1$ and $C(m)$ a constant which may be estimated by $C(m) \leq \Lambda_m m^2$, cf. [BS01]. Affine transformation then yields

$$\|(u - \mathcal{I}_m u)'\|_{\infty,I_j} \leq 2^{-(r-1)}|I_j|^{r-1}\left(\frac{1}{(r-1)!} + \frac{1}{r!}\Lambda_m m^2\right)\|u^{(r)}\|_{\infty,I_j} \tag{29}$$

for $I_j := [a_j, b_j] \subset \mathbb{R}$.
Furthermore, we need an estimate for the norms of the derivatives of algebraic polynomials known as Markov's Theorem [DL93, Theorem 1.4]:

$$\|p'\|_{\infty,I_j} \leq m^2 2|I_j|^{-1}\|p\|_{\infty,I_j} \quad \text{for } p \in \mathcal{P}_m. \tag{30}$$

This estimate cannot be improved; in particular, it is sharp for the Chebyshev polynomials.

To begin with, we state a result concerning the first derivatives of $\kappa - \mathcal{I}_m^{(\sigma,\tau)}\kappa$.

Lemma 3.5. *Provided* $\operatorname{diam}(B_\sigma \times B_\tau) \leq \eta \operatorname{dist}(B_\sigma, B_\tau)$ *with* $\eta > 0$*, an asymptotically smooth* $\kappa \in \mathcal{C}^{m+2}(B)$ *on* $B := B_\sigma \times B_\tau$ *satisfies*

$$\|\partial_\alpha(\kappa - \mathcal{I}_m^{(\sigma,\tau)}\kappa)\|_{\infty,B} \leq c_1c_2 2^{-m}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1} \quad \textit{for } 1 \leq \alpha \leq 2d, \tag{31}$$

where c_1 is a numerical constant depending on d, B, and κ and $c_2 = (\Lambda_m m^2+m+1)\,\Lambda_m^{2d-1}$ grows polynomially in m.

Proof. As in [BG04, BH02], we write

$$\|\partial_\alpha(\kappa - \mathcal{I}_m^{(\sigma,\tau)}\kappa)\|_{\infty,B} \le \sum_{j=1}^{2d} \Big\| \partial_\alpha \underbrace{\Big(\bigotimes_{k=1}^{j-1} \mathcal{I}_m^{I_k} \otimes (Id - \mathcal{I}_m^{I_j}) \otimes \bigotimes_{k=j+1}^{2d} Id \Big)\kappa}_{=:\mathcal{J}_j\kappa} \Big\|_{\infty,B}; \tag{32}$$

to obtain estimates for $\|\partial_\alpha \mathcal{J}_j\kappa\|_{\infty,B}$, we have to consider three cases.

• First, for $j < \alpha$, an explicit computation gives $\partial_\alpha \mathcal{J}_j\kappa = \mathcal{J}_j(\partial_\alpha\kappa)$, whence

$$\begin{aligned}\|\partial_\alpha \mathcal{J}_j\kappa\|_{\infty,B} &\le \Lambda_m^{j-1} \Big\| \Big(\bigotimes_{k=1}^{j-1} Id \otimes (Id - \mathcal{I}_m^{I_j}) \otimes \bigotimes_{k=j+1}^{2d} Id \Big) \partial_\alpha\kappa \Big\|_{\infty,B} \\ &\le \Lambda_m^{j-1} \frac{4^{-m}}{2(m+1)!} |I_j|^{m+1} \|\partial_j^{m+1}\partial_\alpha\kappa\|_{\infty,B};\end{aligned}$$

here we have used (19) on $\mathcal{I}_m^{I_k}$, $k < j$, and applied the estimate in (27) to I_j. By exploiting the asymptotic smoothness of κ to estimate $\|\partial_j^{m+1}\partial_\alpha\kappa\|_{\infty,B} \le C_{\mathrm{asm}}\big(c_{\mathrm{asm}}\,\mathrm{dist}(B_\sigma, B_\tau)\big)^{-(m+2+s)}(m+2)!$, $|I_j| \le \mathrm{diam}(B_\sigma \times B_\tau)$, and $\mathrm{diam}(B_\sigma \times B_\tau) \le \eta\,\mathrm{dist}(B_\sigma, B_\tau)$, we get

$$\|\partial_\alpha \mathcal{J}_j\kappa\|_{\infty,B} \le C_{\mathrm{asm}} c_{\mathrm{asm}}^{-(s+1)} \mathrm{dist}(B_\sigma, B_\tau)^{-(s+1)} c_{\mathrm{asm}}^{-(m+1)} \eta^{m+1} \frac{m+2}{2} 4^{-m} \Lambda_m^{j-1}. \tag{33}$$

• Second, for $j = \alpha$, it follows from (29) with $r = m+1$ that

$$\begin{aligned}\|\partial_\alpha \mathcal{J}_\alpha\kappa\|_{\infty,B} &\le \Lambda_m^{\alpha-1} \Big\| \Big(\bigotimes_{k=1}^{\alpha-1} Id \otimes \partial_\alpha(Id - \mathcal{I}_m^{I_\alpha}) \otimes \bigotimes_{k=\alpha+1}^{2d} Id \Big)\kappa \Big\|_{\infty,B} \\ &\le \Lambda_m^{\alpha-1} 2^{-m} |I_\alpha|^m \Big(\frac{1}{m!} + \frac{m^2}{(m+1)!}\Lambda_m \Big) \|\partial_\alpha^{m+1}\kappa\|_{\infty,B}.\end{aligned}$$

As before, we now obtain

$$\begin{aligned}\|\partial_\alpha \mathcal{J}_\alpha\kappa\|_{\infty,B} \le C_{\mathrm{asm}} c_{\mathrm{asm}}^{-s} \mathrm{dist}(B_\sigma, B_\tau)^{-s} c_{\mathrm{asm}}^{-(m+1)} \eta^{m+1} \\ \times (\Lambda_m m^2 + m + 1) 2^{-m} |I_\alpha|^{-1} \Lambda_m^{\alpha-1}.\end{aligned} \tag{34}$$

• Third, the case $j > \alpha$ is treated by applying Markov's Theorem (30) to $\mathcal{I}_m^{I_\alpha}$ and by using the estimate in (27), whence

$$\begin{aligned}\|\partial_\alpha \mathcal{J}_j\kappa\|_{\infty,B} &\le \Lambda_m^{j-1} m^2 2 |I_\alpha|^{-1} \Big\| \Big(\bigotimes_{k=1}^{j-1} Id \otimes (Id - \mathcal{I}_m^{I_j}) \otimes \bigotimes_{k=j+1}^{2d} Id \Big)\kappa \Big\|_{\infty,B} \\ &\le \Lambda_m^{j-1} m^2 \frac{4^{-m}}{(m+1)!} |I_\alpha|^{-1} |I_j|^{m+1} \|\partial_j^{m+1}\kappa\|_{\infty,B}\end{aligned}$$

and

$$\|\partial_\alpha \mathcal{J}_j\kappa\|_{\infty,B} \le C_{\mathrm{asm}} c_{\mathrm{asm}}^{-s} \mathrm{dist}(B_\sigma, B_\tau)^{-s} c_{\mathrm{asm}}^{-(m+1)} \eta^{m+1} m^2 4^{-m} |I_\alpha|^{-1} \Lambda_m^{j-1}. \tag{35}$$

Collecting the estimates in (33), (34), and (35), we finally have

$$\begin{aligned}
&\|\partial_\alpha(\kappa - \mathcal{I}_m^{(\sigma,\tau)}\kappa)\|_{\infty,B} \\
&\leq C_{\mathrm{asm}}\mathrm{dist}(B_\sigma,B_\tau)^{-s}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}c_{\mathrm{asm}}^{-s}\bigg(c_{\mathrm{asm}}^{-1}\frac{m+2}{2}4^{-m}\mathrm{dist}(B_\sigma,B_\tau)^{-1}\sum_{j<\alpha}\Lambda_m^{j-1} \\
&\qquad + (\Lambda_m m^2 + m + 1)2^{-m}|I_\alpha|^{-1}\Lambda_m^{\alpha-1} + m^2 4^{-m}|I_\alpha|^{-1}\sum_{j>\alpha}\Lambda_m^{j-1}\bigg) \\
&\leq c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}(\Lambda_m m^2 + m + 1)2^{-m}\Lambda_m^{2d-1}C_{\mathrm{asm}}c_{\mathrm{asm}}^{-s}\mathrm{dist}(B_\sigma,B_\tau)^{-s} \\
&\qquad \times\Big((\alpha-1)c_{\mathrm{asm}}^{-1}\mathrm{dist}(B_\sigma,B_\tau)^{-1} + \big(2d-(\alpha-1)\big)\big(\min_{j=1}^{2d}|I_j|\big)^{-1}\Big),
\end{aligned}$$

which gives the desired result. □

For the second derivatives of $\kappa - \mathcal{I}_m^{(\sigma,\tau)}\kappa$, we obtain in a similar fashion:

Lemma 3.6. *Under the assumptions of Lemma 3.5, we have*

$$\|\partial_\alpha\partial_\beta(\kappa - \mathcal{I}_m^{(\sigma,\tau)}\kappa)\|_{\infty,B} \leq c_1 c_2 2^{-(m-1)}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1} \quad \text{for } 1\leq\alpha,\beta\leq 2d, \tag{36}$$

with a constant c_1 depending on d, B, and κ and $c_2 = (\Lambda_m m^2 + m + 1)m^2\Lambda_m^{2d-1}$.

Proof. Without loss of generality, we assume $1\leq\alpha\leq\beta\leq 2d$ throughout; similar reasoning as in the proof of Lemma 3.5, with $\mathcal{J}_j\kappa$ defined as in (32), then implies the following cases:

- $j<\alpha\leq\beta$: with $\partial_\alpha\partial_\beta\mathcal{J}_j\kappa = \mathcal{J}_j(\partial_\alpha\partial_\beta\kappa)$, one has as in (33)

$$\begin{aligned}
\|\partial_\alpha\partial_\beta\mathcal{J}_j\kappa\|_{\infty,B} &\leq \Lambda_m^{j-1}\frac{4^{-m}}{2(m+1)!}|I_j|^{m+1}\|\partial_j^{m+1}\partial_\alpha\partial_\beta\kappa\|_{\infty,B} \\
&\leq \Lambda_m^{j-1}C_{\mathrm{asm}}c_{\mathrm{asm}}^{-(s+2)}\mathrm{dist}(B_\sigma,B_\tau)^{-(s+2)}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}4^{-m}\frac{(m+2)(m+3)}{2};
\end{aligned}$$

- $j=\alpha<\beta$: from $\partial_\alpha\partial_\beta\mathcal{J}_\alpha\kappa = \partial_\alpha\mathcal{J}_\alpha(\partial_\beta\kappa)$ and (29) for $r=m+1$ it follows that

$$\begin{aligned}
\|\partial_\alpha\partial_\beta\mathcal{J}_\alpha\kappa\|_{\infty,B} &\leq \Lambda_m^{\alpha-1}2^{-m}|I_\alpha|^m\bigg(\frac{1}{m!}+\frac{m^2}{(m+1)!}\Lambda_m\bigg)\|\partial_j^{m+1}\partial_\beta\kappa\|_{\infty,B} \\
&\leq \Lambda_m^{\alpha-1}C_{\mathrm{asm}}c_{\mathrm{asm}}^{-(s+1)}\mathrm{dist}(B_\sigma,B_\tau)^{-(s+1)}|I_\alpha|^{-1}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}2^{-m} \\
&\qquad\qquad \times(m+2)(1+m+m^2\Lambda_m),
\end{aligned}$$

  cf. (34);

- $\alpha<j<\beta$: as $\partial_\alpha\partial_\beta\mathcal{J}_j\kappa = \partial_\alpha\mathcal{J}_j(\partial_\beta\kappa)$, we obtain with (30)

$$\begin{aligned}
\|\partial_\alpha\partial_\beta\mathcal{J}_j\kappa\|_{\infty,B} &\leq \Lambda_m^{j-1}m^2 2|I_\alpha|^{-1}\frac{4^{-m}}{2(m+1)!}|I_j|^{m+1}\|\partial_j^{m+1}\partial_\beta\kappa\|_{\infty,B} \\
&\leq \Lambda_m^{j-1}C_{\mathrm{asm}}c_{\mathrm{asm}}^{-(s+1)}\mathrm{dist}(B_\sigma,B_\tau)^{-(s+1)}|I_\alpha|^{-1}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}4^{-m}m^2(m+2),
\end{aligned}$$

  see (35);

- $j=\alpha=\beta$: to obtain an estimate for $\partial_j^2\mathcal{J}_j\kappa$, we proceed as in [BS01, Theorem 3.3.1]: given $u\in\mathcal{C}^{m+1}[-1,1]$ and $1\leq r\leq m+1$, let R_r be the remainder term of the Taylor series expansion of degree $r-1$ of u about $t=0$. Using $u-\mathcal{I}_m u = R_r - \mathcal{I}_m R_r$,

$\|R_r\|_{\infty,[-1,1]} \leq 1/r!\|u^{(r)}\|_{\infty,[-1,1]}$, and $\|R_r''\|_{\infty,[-1,1]} \leq 1/(r-2)!\|u^{(r)}\|_{\infty,[-1,1]}$, we have by the triangle inequality and with (30)

$$\begin{aligned}\|(u-\mathcal{I}_m u)''\|_{\infty,[-1,1]} &\leq \|R_r''\|_{\infty,[-1,1]} + \|(\mathcal{I}_m R_r)''\|_{\infty,[-1,1]} \\ &\leq \frac{1}{(r-2)!}\|u^{(r)}\|_{\infty,[-1,1]} + m^2(m-1)^2\Lambda_m\frac{1}{r!}\|u^{(r)}\|_{\infty,[-1,1]}.\end{aligned}$$

By affine transformation it follows for $u \in \mathcal{C}^{m+1}(I_\alpha)$ and $r = m+1$ that

$$\begin{aligned}&\|(u-\mathcal{I}_m u)''\|_{\infty,I_\alpha} \\ &\qquad \leq 2^{-(m-1)}|I_\alpha|^{m-1}\bigg(\frac{1}{(m-1)!} + \frac{1}{(m+1)!}(m-1)^2m^2\Lambda_m\bigg)\|u^{(m+1)}\|_{\infty,I_\alpha},\end{aligned}$$

whence

$$\begin{aligned}\|\partial_\alpha^2\mathcal{J}_\alpha\kappa\|_{\infty,B} &\leq \Lambda_m^{\alpha-1}\bigg\|\bigg(\bigotimes_{k=1}^{\alpha-1} Id \otimes \partial_\alpha^2(Id-\mathcal{I}_m^{I_\alpha}) \otimes \bigotimes_{k=\alpha+1}^{2d} Id\bigg)\kappa\bigg\|_{\infty,B} \\ &\leq \Lambda_m^{\alpha-1}2^{-(m-1)}|I_\alpha|^{m-1}\bigg(\frac{1}{(m-1)!} + \frac{1}{(m+1)!}(m-1)^2m^2\Lambda_m\bigg)\|\partial_\alpha^{m+1}\kappa\|_{\infty,B} \\ &\leq \Lambda_m^{\alpha-1}C_{\mathrm{asm}}c_{\mathrm{asm}}^{-s}\mathrm{dist}(B_\sigma,B_\tau)^{-s}|I_\alpha|^{-2}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}2^{-(m-1)} \\ &\qquad\qquad \times \big(m(m+1)+(m-1)^2m^2\Lambda_m\big);\end{aligned}$$

- $\alpha < j = \beta$: with $\partial_\alpha\partial_\beta\mathcal{J}_\beta\kappa = \partial_\alpha(\partial_\beta\mathcal{J}_\beta\kappa)$, (29) and (30) give

$$\begin{aligned}\|\partial_\alpha\partial_\beta\mathcal{J}_\beta\kappa\|_{\infty,B} &\leq \Lambda_m^{\beta-1}m^2 2|I_\alpha|^{-1}2^{-m}|I_\beta|^m\bigg(\frac{1}{m!} + \frac{m^2}{(m+1)!}\Lambda_m\bigg)\|\partial_\beta^{m+1}\kappa\|_{\infty,B} \\ &\leq \Lambda_m^{\beta-1}C_{\mathrm{asm}}c_{\mathrm{asm}}^{-s}\mathrm{dist}(B_\sigma,B_\tau)^{-s}|I_\alpha|^{-1}|I_\beta|^{-1}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}2^{-(m-1)} \\ &\qquad\qquad \times m^2\big(1+m+m^2\Lambda_m\big);\end{aligned}$$

- $\alpha \leq \beta < j$: applying Markov's Theorem (30) twice yields

$$\begin{aligned}\|\partial_\alpha\partial_\beta\mathcal{J}_j\kappa\|_{\infty,B} &\leq \Lambda_m^{j-1}m^2 2|I_\alpha|^{-1}m^2 2|I_\beta|^{-1}\frac{4^{-m}}{2(m+1)!}|I_j|^{m+1}\|\partial_j^{m+1}\partial_\alpha\kappa\|_{\infty,B} \\ &\leq \Lambda_m^{j-1}C_{\mathrm{asm}}c_{\mathrm{asm}}^{-s}\mathrm{dist}(B_\sigma,B_\tau)^{-s}|I_\alpha|^{-1}|I_\beta|^{-1}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1}4^{-(m-1)}\frac{m^4}{2}.\end{aligned}$$

Let us first consider $\alpha < \beta$: by collecting the above estimates, we obtain as in Lemma 3.5

$$\begin{aligned}&\|\partial_\alpha\partial_\beta(\kappa-\mathcal{I}_m^{(\sigma,\tau)}\kappa)\|_{\infty,B} \\ &\quad\leq C_{\mathrm{asm}}c_{\mathrm{asm}}^{-s}\mathrm{dist}(B_\sigma,B_\tau)^{-s}(\Lambda_m m^2+m+1)m^2 2^{-(m-1)}\Lambda_m^{2d-1}c_{\mathrm{asm}}^{-(m+1)}\eta^{m+1} \\ &\qquad\times\Big((\alpha-1)c_{\mathrm{asm}}^{-2}\mathrm{dist}(B_\sigma,B_\tau)^{-2} + (\beta-\alpha)c_{\mathrm{asm}}^{-1}\mathrm{dist}(B_\sigma,B_\tau)^{-1}\big(\min_{j=1}^{2d}|I_j|\big)^{-1} \\ &\qquad\qquad + \big(2d-(\beta-1)\big)\big(\min_{j=1}^{2d}|I_j|\big)^{-2}\Big),\end{aligned}$$

from which the result follows. An analogous computation for $\alpha = \beta$ shows that the estimate then still holds, with the same constants c_1 and c_2. This concludes the proof. □

***Remark* 7.** Lemmas 3.5 and 3.6 ensure exponential convergence with respect to m provided at least $2c_{\mathrm{asm}} > \eta$. For κ the Newtonian kernel G, this implies exponential convergence for $\eta \in (0, 2)$, cf. Remark 4.

4. $\mathcal{H}^2$-Matrix Techniques

In this section, we motivate the concept of hierarchical matrices and give the corresponding definitions.

4.1. **Motivation.** We consider a bilinear form $a(\cdot,\cdot)$ on $L^2(\Omega)$ given by a double integration with an asymptotically smooth kernel κ,

$$a(u,v) := \int_\Omega \int_\Omega u(x)\kappa(x,y)v(y)\,dy\,dx \quad \text{for } u, v \in L^2(\Omega), \tag{37}$$

with $\Omega \subset \mathbb{R}^d$ a bounded domain. In the following we require a partition $\mathcal{T}$ of Ω and a *block partitioning* $\mathbb{P}$ of $\mathcal{T} \times \mathcal{T}$. For $\eta > 0$ fixed, a block $(\sigma, \tau) \in \mathbb{P}$ is then called *admissible* provided

$$\operatorname{diam}(B_\sigma \times B_\tau) \le \eta \operatorname{dist}(B_\sigma, B_\tau), \tag{38}$$

where B_σ and B_τ denote axially parallel boxes in $\mathbb{R}^d$ of minimal size containing $\cup\sigma$ and $\cup\tau$, respectively; otherwise (σ, τ) is called *inadmissible*. Here $\cup\sigma := \big\{x \in T \,\big|\, T \in \sigma \subseteq \mathcal{T}\big\}$ (resp. $\cup\tau := \big\{y \in T \,\big|\, T \in \tau \subseteq \mathcal{T}\big\}$) denotes the union of all elements $T \in \mathcal{T}$ contained in σ (resp. in τ). $\mathbb{P}$ is thus split into two subsets $\mathbb{P}_{\mathrm{far}}$ and $\mathbb{P}_{\mathrm{near}}$: the subset of all admissible blocks is called *far field* and denoted by $\mathbb{P}_{\mathrm{far}}$; the inadmissible blocks are collected in the *near field* $\mathbb{P}_{\mathrm{near}} := \mathbb{P}\backslash\mathbb{P}_{\mathrm{far}}$.

The approximate bilinear form $\widetilde{a}(\cdot,\cdot)$ is obtained by replacing the kernel function on admissible blocks by an approximate but degenerate kernel obtained by interpolation. More precisely,

$$\begin{aligned}\widetilde{a}(u,v) := &\sum_{(\sigma,\tau)\in\mathbb{P}_{\mathrm{near}}} \int_{\cup\sigma}\int_{\cup\tau} u(x)\kappa(x,y)v(y)\,dy\,dx \\ &+ \sum_{(\sigma,\tau)\in\mathbb{P}_{\mathrm{far}}} \int_{\cup\sigma}\int_{\cup\tau} u(x)\big(\mathcal{I}_m^{(\sigma,\tau)}\kappa(x,y)\big)v(y)\,dy\,dx,\end{aligned} \tag{39}$$

where $\mathcal{I}_m^{(\sigma,\tau)}$ denotes the tensor interpolation operator with respect to the bounding box $B := B_\sigma \times B_\tau$.
For each $\tau \subseteq \mathcal{T}$ with corresponding bounding box B_τ, define a family $(x_j^\tau)_{j=0}^{M_\tau}$ of interpolation points plus the associated tensor Lagrange polynomials $(\mathcal{L}_j^\tau)_{j=0}^{M_\tau}$, where $M_\tau \in \mathbb{N}$. Given

$$\mathcal{I}_m^{(\sigma,\tau)}\kappa(x,y) = \sum_{m_1=0}^{M_\sigma}\sum_{m_2=0}^{M_\tau} \kappa\big(x_{m_1}^\sigma, x_{m_2}^\tau\big)\mathcal{L}_{m_1}^\sigma(x)\mathcal{L}_{m_2}^\tau(y) \quad \text{for } (x,y) \in \cup\sigma \times \cup\tau,$$

the second term from (39) can then be written as

$$\sum_{(\sigma,\tau)\in\mathbb{P}_{\mathrm{far}}}\sum_{m_1=0}^{M_\sigma}\sum_{m_2=0}^{M_\tau} \underbrace{\kappa\big(x_{m_1}^\sigma, x_{m_2}^\tau\big)}_{=:S_{m_1m_2}^{\sigma\tau}} \underbrace{\int_{\cup\sigma} u(x)\mathcal{L}_{m_1}^\sigma(x)\,dx}_{=:V_{m_1}^\sigma(u)} \underbrace{\int_{\cup\tau} v(y)\mathcal{L}_{m_2}^\tau(y)\,dy}_{=:V_{m_2}^\tau(v)}. \tag{40}$$

The advantage of this new representation becomes obvious if we discretize $a(\cdot,\cdot)$: consider the basis $\{\varphi_1,\dots,\varphi_N\}$ of the space $\mathcal{P}_0(\mathcal{T})$ of piecewise constant functions on $\mathcal{T}$ given by $\varphi_j := \chi_{T_j}$, where χ_{T_j} is the characteristic function on $T_j \in \mathcal{T}$ and $N = |\mathcal{T}|$, and define $A \in \mathbb{R}^{N\times N}$ by $A_{jk} := a(\varphi_j,\varphi_k)$ and the approximate matrix $\widetilde{A} \in \mathbb{R}^{N\times N}$ by $\widetilde{A}_{jk} := \widetilde{a}(\varphi_j,\varphi_k)$ for all $1 \le j,k \le N$. On inadmissible blocks $(\sigma,\tau) \in \mathbb{P}_{\text{near}}$, we then simply have $\widetilde{A}|_{\sigma\times\tau} = A|_{\sigma\times\tau}$, whereas for $(\sigma,\tau) \in \mathbb{P}_{\text{far}}$, $\widetilde{A}$ is given by

$$\widetilde{A}|_{\sigma\times\tau} = V^\sigma S^{\sigma\tau} V^{\tau T} \approx A|_{\sigma\times\tau}, \tag{41}$$

with $V^\sigma_{jm_1} := V^\sigma_{m_1}(\varphi_j)$ and $V^\tau_{km_2} := V^\tau_{m_2}(\varphi_k)$ as defined in (40).

4.2. Block Partitioning. A simple method to find a hierarchical partition $\mathbb{P}$ of $\mathcal{T}\times\mathcal{T}$ is to first construct a *cluster tree* from $\mathcal{T}$ by *binary space partitioning*: one starts with the root cluster containing all $T \in \mathcal{T}$, splits it into two son clusters and repeats the procedure recursively until each cluster contains less than a given number of elements $C_{\ell\text{f}}$. Geometrically speaking, for every $T \in \mathcal{T}$, one chooses a coordinate axis and splits the set along this axis.

One can then use the admissibility condition (38) in combination with the cluster tree structure to construct $\mathbb{P}$: starting with $\mathcal{T}\times\mathcal{T}$, one splits each pair of clusters as long as it is not admissible. This gives a $\mathbb{P}$ satisfying $\mathcal{T}\times\mathcal{T} = \bigcup_{(\sigma,\tau)\in\mathbb{P}} \sigma\times\tau$. Clearly, a pair (σ,τ) can only appear in $\mathbb{P}$ if it is admissible or if either σ or τ is a leaf. The above procedure can easily be formalized, see e.g. [BGH03, BH02, HKS00] for formal definitions and algorithms.

4.3. $\mathcal{H}$-Matrices vs. $\mathcal{H}^2$-Matrices. Based on the concepts of cluster tree and block partitioning, the matrix approximation approach outlined in Section **4.1** can be generalized by introducing a class of data-sparse matrices, the so-called $\mathcal{H}$-matrices. Given a block partitioning $\mathbb{P}$ of $\mathcal{T}\times\mathcal{T}$ and some $k \in \mathbb{N}$, a matrix $A \in \mathbb{R}^{N\times N}$ is called *$\mathcal{H}$-matrix of rank k* provided $\operatorname{rank}(A|_{\sigma\times\tau}) \le k$ for each $(\sigma,\tau) \in \mathbb{P}$. Moreover, if a factorization of the form (41) holds for a family $V = (V^\tau)_{\tau\subseteq\mathcal{T}}$ and some *multiplication matrices* $S^{\sigma\tau}$, A is called *uniform $\mathcal{H}$-matrix with respect to the cluster basis V*, cf. [BGH03, HKS00].

Additional structure can be gained by considering uniform $\mathcal{H}$-matrices for which the corresponding cluster bases are *nested*: if the space $\mathcal{Q}_m$ defined in (21) is used for interpolation on *all* clusters, polynomials corresponding to father clusters can be expressed *exactly* in terms of polynomials corresponding to son clusters. For $\tau \subseteq \mathcal{T}$ and $\tau' \in \operatorname{sons}(\tau)$, we have

$$\mathcal{L}^\tau_j(x) = \sum_{m=0}^{M} \mathcal{L}^\tau_j\big(x^{\tau'}_m\big)\mathcal{L}^{\tau'}_m(x). \tag{42}$$

Defining the *transfer matrix* $B^{\tau'\tau} \in \mathbb{R}^{M\times M}$ by $B^{\tau'\tau}_{mj} := \mathcal{L}^\tau_j\big(x^{\tau'}_m\big)$, we obtain

$$\begin{aligned} V^\tau_{ij} = \int_{\cup\tau} \chi_{T_j}\mathcal{L}^\tau_j(x)dx &= \sum_{\tau'\in\operatorname{sons}(\tau)} \sum_{m=0}^{M} \mathcal{L}^\tau_j\big(x^{\tau'}_m\big) \int_{\cup\tau'} \chi_{T_j}\mathcal{L}^{\tau'}_m(x)dx \\ &= \sum_{\tau'\in\operatorname{sons}(\tau)} \sum_{m=0}^{M} B^{\tau'\tau}_{mj} V^{\tau'}_{im} \end{aligned} \tag{43}$$

for all $i \in \tau' \subset \tau$, i.e., $V^\tau|_{\tau'} = V^{\tau'} B^{\tau' \tau}$. A is called *$\mathcal{H}^2$-matrix with respect to V* if it is a uniform $\mathcal{H}$-matrix with respect to V and if V is nested.

***Remark* 8.** For $\mathcal{H}^2$-matrices, one only has to store the multiplication matrices $S^{\sigma\tau}$ for all admissible blocks $(\sigma, \tau) \in \mathbb{P}_{\text{far}}$, the cluster matrices V^τ for all leaves $\tau \subseteq \mathcal{T}$, the transfer matrices $B^{\tau' \tau}$ for all father-son pairs, and $A|_{\sigma \times \tau}$ on all non-admissible blocks $(\sigma, \tau) \in \mathbb{P}_{\text{near}}$.

***Remark* 9.** To evaluate (39), fast and efficient algorithms for matrix-vector multiplication are required. We assume the underlying partitioning $\mathbb{P}$ to be *sparse* in the sense of [Gra01], with some sparsity constant C_{sp}. Given an $\mathcal{H}^2$-matrix $\widetilde{A} \in \mathbb{R}^{N \times N}$ and $x, y \in \mathbb{R}^N$, it can then be shown that the computation of $y = \widetilde{A}x$ needs only $\mathcal{O}(Nm^d)$ operations to complete. Similarly, both the number of operations required to build an $\mathcal{H}^2$-matrix approximation and the amount of storage needed to store it are of order $\mathcal{O}(Nm^d)$, cf. [BGH03, Gie01, HKS00]. Note that C_{sp} enters all the above complexity estimates; for $C_{\text{sp}} \to \infty$, the complexity of the problem will become unbounded, as well. Estimates on C_{sp} can be found in [GH03].

4.4. Global Error Analysis for the Approximate Bilinear Form. By definition of the admissibility condition (38), one can apply the approximation results of Section **3.4** on each admissible block $(\sigma, \tau) \in \mathbb{P}_{\text{far}}$. Indeed, given some constant c_3, Lemma 3.4 shows that we can choose an approximation order $m \in \mathbb{N}$ so that

$$\|\kappa - \mathcal{I}_m^{(\sigma,\tau)} \kappa\|_{L^\infty(B_\sigma \times B_\tau)} \leq c_3$$

for all $(\sigma, \tau) \in \mathbb{P}_{\text{far}}$.

Theorem 4.1. *Under the above assumptions, we have*

$$|a(u,v) - \widetilde{a}(u,v)| \leq c_3 \|u\|_{L^1(\Omega)} \|v\|_{L^1(\Omega)} \leq c_3 |\Omega| \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \tag{44}$$

for all $u, v \in L^2(\Omega)$.

Proof. For almost all $(x, y) \in \Omega \times \Omega$, we define an integral kernel $\widetilde{\kappa}(x, y)$ as follows. Let $\mathcal{S} := \bigcup \{\partial T \,|\, T \in \mathcal{T}\}$ denote the skeleton of the partition. Note that $\mathcal{S} \subset \mathbb{R}^d$ is a set of measure zero. For $x, y \in \Omega \backslash \mathcal{S}$, there are *unique* elements $T_x, T_y \in \mathcal{T}$ satisfying $x \in T_x$ and $y \in T_y$, respectively. Since $\mathbb{P}$ is a partition of $\mathcal{T} \times \mathcal{T}$, there is a unique block $(\sigma, \tau) \in \mathbb{P}$ with $(T_x, T_y) \in \sigma \times \tau$. Consequently, we may define

$$\widetilde{\kappa}(x,y) := \begin{cases} \kappa(x,y) & \text{if } (\sigma,\tau) \text{ is not admissible,} \\ \mathcal{I}_m^{(\sigma,\tau)} \kappa(x,y) & \text{else.} \end{cases}$$

Since $a(\cdot,\cdot)$ and $\widetilde{a}(\cdot,\cdot)$ differ only on the far-field blocks, we have

$$|a(u,v) - \widetilde{a}(u,v)| = \left| \int_\Omega \int_\Omega u(x) \big(\kappa(x,y) - \widetilde{\kappa}(x,y)\big) v(y) \, dy \, dx \right|,$$

and a Hölder inequality yields

$$|a(u,v) - \widetilde{a}(u,v)| \leq \|\kappa - \widetilde{\kappa}\|_{L^\infty(\Omega \times \Omega)} \|u\|_{L^1(\Omega)} \|v\|_{L^1(\Omega)},$$

since $u(x)$ and $v(y)$ decouple on $\Omega \times \Omega$. Using $\|u\|_{L^1(\Omega)} \leq |\Omega|^{1/2} \|u\|_{L^2(\Omega)}$, we obtain the desired estimate. □

4.5. Global Error Analysis for the Corresponding Matrix. Let $U_h \leq L^2(\Omega)$ be a finite-dimensional space, and define the stiffness matrix $A \in \mathbb{R}^{N\times N}$ and its $\mathcal{H}^2$-approximation $\widetilde{A} \in \mathbb{R}^{N\times N}$ by $A_{jk} := a(u_j, u_k)$ and $\widetilde{A}_{jk} := \widetilde{a}(u_j, u_k)$, respectively, for a fixed basis $\{u_1, \dots, u_N\}$ of U_h. Then, there holds

Corollary 4.2. *The approximation error for the stiffness matrix A is bounded by*

$$\|A - \widetilde{A}\|_F \leq N c_3 \big(\max\nolimits_{j=1}^N \|u_j\|_{L^1(\Omega)}\big)^2. \tag{45}$$

In particular, the error decreases exponentially with the approximation order m. Provided A is a regular matrix, $\widetilde{A}$ also is regular for large approximation orders m.

Proof. The first assertion follows by applying Theorem 4.1 to each entry of A. Regularity is immediate, as the regular matrices $GL(N)$ form an open subset of $\mathbb{R}^{N\times N}$. □

5. Galerkin Discretization of the Potential Equation

In this section we provide two different $\mathcal{H}^2$-matrix approaches to obtain a reasonable *data sparse* approximation of the stiffness matrix

$$\mathbf{A} \in \mathbb{R}^{dN\times dN} \quad \text{with} \quad A_{jk} := a(\varphi_j, \varphi_k) \tag{46}$$

for a fixed basis $\{\varphi_1, \dots, \varphi_{dN}\}$ of $\mathcal{P}_0(\mathcal{T})^d$, where the bilinear form $a(\cdot,\cdot)$ is defined as in (15). We recall a result from [Pra04].

Theorem 5.1. *For bounded Lipschitz domains $\omega, \widetilde{\omega} \subset \mathbb{R}^d$ and vectors $\mathbf{m}, \widetilde{\mathbf{m}} \in \mathbb{R}^d$, we have*

$$a(\chi_\omega \mathbf{m}, \chi_{\widetilde{\omega}} \widetilde{\mathbf{m}}) = -\int_{\partial\omega}\int_{\partial\widetilde{\omega}} G(x-y)\big(\boldsymbol{\nu}(x)\cdot\mathbf{m}\big)\big(\widetilde{\boldsymbol{\nu}}(y)\cdot\widetilde{\mathbf{m}}\big)\, ds_y\, ds_x, \tag{47}$$

where $\boldsymbol{\nu}$ and $\widetilde{\boldsymbol{\nu}}$ denote the outer normal vectors on $\partial\omega$ and $\partial\widetilde{\omega}$, respectively. Furthermore, we have the symmetry properties

$$a(\chi_\omega \mathbf{m}, \chi_{\widetilde{\omega}} \widetilde{\mathbf{m}}) = a(\chi_{\widetilde{\omega}} \widetilde{\mathbf{m}}, \chi_\omega \mathbf{m}) = a(\chi_\omega \widetilde{\mathbf{m}}, \chi_{\widetilde{\omega}} \mathbf{m}), \tag{48}$$

and in the case $\operatorname{dist}(\omega, \widetilde{\omega}) > 0$ there holds

$$a(\chi_\omega \mathbf{m}, \chi_{\widetilde{\omega}} \widetilde{\mathbf{m}}) = \int_\omega \int_{\widetilde{\omega}} \mathbf{m} \cdot H_G(x-y) \widetilde{\mathbf{m}}\, dy\, dx \tag{49}$$

with the Hessian H_G of the Newtonian kernel G. □

Now, a reasonable choice for a basis of $\mathcal{P}_0(\mathcal{T})^d$ is

$$\varphi_j := \chi_{T_j}\mathbf{e}_1, \quad \varphi_{j+N} := \chi_{T_j}\mathbf{e}_2 \quad \text{etc.} \quad \text{for } 1 \leq j \leq N, \tag{50}$$

as is shown in the following. This basis gives rise to the definition of the matrices

$$\mathbf{A}^{\alpha\beta} \in \mathbb{R}^{N\times N}_{sym} \quad \text{for fixed } 1 \leq \alpha, \beta \leq d, \quad A^{\alpha\beta}_{jk} := a(\chi_{T_j}\mathbf{e}_\alpha, \chi_{T_k}\mathbf{e}_\beta), \tag{51}$$

where the symmetry of $\mathbf{A}^{\alpha\beta}$ (i.e., an additional symmetry of $\mathbf{A}$) follows from (48). Note that – again by equation (48) – we have $\mathbf{A}^{\alpha\beta} = \mathbf{A}^{\beta\alpha}$. Therefore, $\mathbf{A}$ is a symmetric $d \times d$ block matrix with symmetric blocks $\mathbf{A}^{\alpha\beta}$ of dimension $N \times N$,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{12} & \mathbf{A}^{22} \end{pmatrix} \quad \text{for } d = 2 \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} & \mathbf{A}^{13} \\ \mathbf{A}^{12} & \mathbf{A}^{22} & \mathbf{A}^{23} \\ \mathbf{A}^{13} & \mathbf{A}^{23} & \mathbf{A}^{33} \end{pmatrix} \quad \text{for } d = 3, \tag{52}$$

respectively. The idea is to approximate each block $\mathbf{A}^{\alpha\beta}$ by an appropriate $\mathcal{H}^2$-matrix.

5.1. **Approximation of A on Far Field Blocks via (49).** A direct $\mathcal{H}^2$-matrix approach stems from the far field representation (49) for which the Hessian (i.e., the second derivatives) of the Newtonian kernel enters,

$$\frac{\partial^2 G}{\partial x_\alpha \partial x_\beta}(x) = \frac{1}{\gamma_d}\frac{\delta_{\alpha\beta}|x|^2 - dx_\alpha x_\beta}{|x|^{d+2}} \quad \text{for } x \in \mathbb{R}^d\backslash\{0\}, \tag{53}$$

where $1 \le \alpha, \beta \le d$ and $\delta_{\alpha\beta}$ denotes the Kronecker delta. Obviously, these kernels are asymptotically smooth with singularity order $s = -d$, cf. Remark 4. Let $\mathbb{P}$ be a block partitioning with respect to the given triangulation $\mathcal{T}$. To abbreviate notation, let $\boldsymbol{\nu}^j$ denote the outer normal vector on the boundary ∂T_j of an element T_j, and let $\mathbf{m}_h^j := \mathbf{m}_h|_{T_j} \in \mathbb{R}^d$ be a discrete magnetization $\mathbf{m}_h \in \mathcal{P}_0(\mathcal{T})^d$. In analogy to the previous section and according to (47) and (49), the bilinear form $a(\cdot,\cdot)$ on $\mathcal{P}_0(\mathcal{T})^d$ reads

$$\begin{aligned} a(\mathbf{m}_h, \widetilde{\mathbf{m}}_h) &= -\sum_{j,k=1}^{N} \int_{\partial T_j}\int_{\partial T_k} G(x-y)\big(\boldsymbol{\nu}^j(x)\cdot\mathbf{m}_h^j\big)\big(\boldsymbol{\nu}^k(y)\cdot\widetilde{\mathbf{m}}_h^k\big)\,ds_y\,ds_x \\ &= -\sum_{(\sigma,\tau)\in\mathbb{P}_{\text{near}}}\sum_{T_j\in\sigma}\sum_{T_k\in\tau}\int_{\partial T_j}\int_{\partial T_k} G(x-y)\big(\boldsymbol{\nu}^j(x)\cdot\mathbf{m}_h^j\big)\big(\boldsymbol{\nu}^k(y)\cdot\widetilde{\mathbf{m}}_h^k\big)\,ds_y\,ds_x \\ &\qquad + \sum_{(\sigma,\tau)\in\mathbb{P}_{\text{far}}}\int_{\cup\sigma}\int_{\cup\tau}\mathbf{m}_h(x)\cdot H_G(x-y)\,\widetilde{\mathbf{m}}_h(y)\,dy\,dx. \end{aligned} \tag{54}$$

As in the previous section, we obtain the approximate bilinear form by replacing the exact kernel H_G on far field blocks by tensor interpolation $\mathcal{I}_m^{(\sigma,\tau)}H_G$ which is now understood coefficient-wise [since we are dealing with a matrix kernel $H_G(x-y) \in \mathbb{R}^{d\times d}_{sym}$],

$$\begin{aligned} &\widetilde{a}(\mathbf{m}_h, \widetilde{\mathbf{m}}_h) \\ &:= -\sum_{(\sigma,\tau)\in\mathbb{P}_{\text{near}}}\sum_{T_j\in\sigma}\sum_{T_k\in\tau}\int_{\partial T_j}\int_{\partial T_k} G(x-y)\big(\boldsymbol{\nu}^j(x)\cdot\mathbf{m}_h^j\big)\big(\boldsymbol{\nu}^k(y)\cdot\widetilde{\mathbf{m}}_h^k\big)\,ds_y\,ds_x \\ &\qquad + \sum_{(\sigma,\tau)\in\mathbb{P}_{\text{far}}}\int_{\cup\sigma}\int_{\cup\tau}\mathbf{m}_h(x)\cdot\big(\mathcal{I}_m^{(\sigma,\tau)}H_G(x-y)\big)\,\widetilde{\mathbf{m}}_h(y)\,dy\,dx. \end{aligned}$$

The error analysis is completely straightforward following the arguments of Section **4.4**. We denote by $\kappa_{\alpha\beta}$, $1 \le \alpha, \beta \le d$, the second derivatives of the Newtonian kernel, cf. (53). As before, we assume that the approximation degree $m \in \mathbb{N}$ is large enough so that

$$\|\kappa_{\alpha\beta} - \mathcal{I}_m^{(\sigma,\tau)}\kappa_{\alpha\beta}\|_{L^\infty(B_\sigma\times B_\tau)} \le c_3,$$

with a constant c_3 which is independent of $(\sigma,\tau) \in \mathbb{P}_{\text{far}}$.

Theorem 5.2. *Under the above assumptions, we have*

$$|a(\mathbf{m}_h, \widetilde{\mathbf{m}}_h) - \widetilde{a}(\mathbf{m}_h, \widetilde{\mathbf{m}}_h)| \le c_3 d|\Omega|\|\mathbf{m}_h\|_{L^2(\Omega)}\|\widetilde{\mathbf{m}}_h\|_{L^2(\Omega)} \tag{55}$$

for all $\mathbf{m}_h, \widetilde{\mathbf{m}}_h \in \mathcal{P}_0(\mathcal{T})^d$.

Proof. With $\kappa(x,y) := H_G(x-y)$ and $\widetilde{\kappa}$ as in the proof of Theorem 4.1, we obtain

$$|a(\mathbf{m}_h, \widetilde{\mathbf{m}}_h) - \widetilde{a}(\mathbf{m}_h, \widetilde{\mathbf{m}}_h)| \le |\Omega| \|\kappa - \widetilde{\kappa}\|_{L^\infty(\Omega\times\Omega;\mathbb{R}^{d\times d})} \|\mathbf{m}_h\|_{L^2(\Omega;\mathbb{R}^d)} \|\widetilde{\mathbf{m}}_h\|_{L^2(\Omega;\mathbb{R}^d)},$$

where we consider the usual (Euclidean) operator norm $\|\cdot\|$ on $\mathbb{R}^{d\times d}$. Recalling that the Frobenius norm satisfies $\|A\| \le \|A\|_F := \big(\sum_{j,k=1}^d A_{jk}^2\big)^{1/2}$, it follows that

$$\|\kappa(x,y) - \widetilde{\kappa}(x,y)\| \le \|\kappa(x,y) - \widetilde{\kappa}(x,y)\|_F \le dc_3.$$

This concludes the proof. □

Finally, we explicitly state the approximation $\widetilde{\mathbf{A}}^{\alpha\beta} \in \mathbb{R}^{N\times N}$ to $\mathbf{A}^{\alpha\beta}$ to clarify what has to be implemented. Recall that $\widetilde{\mathbf{A}}^{\alpha\beta} \in \mathbb{R}^{N\times N}$ is defined by $\widetilde{A}_{jk}^{\alpha\beta} = \widetilde{a}(\chi_{T_j}\mathbf{e}_\alpha, \chi_{T_k}\mathbf{e}_\beta)$. The computation of $\widetilde{\mathbf{A}}^{\alpha\beta}$ is performed separately on the admissible and the inadmissible blocks of $\mathbb{P}$.

First, let $(\sigma,\tau) \in \mathbb{P}_{\mathrm{far}}$ be admissible; given the degenerate kernel

(56)

$$\mathcal{I}_m^{(\sigma,\tau)}\kappa_{\alpha\beta}(x,y) = \sum_{m_1=0}^{M}\sum_{m_2=0}^{M} \kappa_{\alpha\beta}\big(x_{m_1}^\sigma, x_{m_2}^\tau\big)\mathcal{L}_{m_1}^\sigma(x)\mathcal{L}_{m_2}^\tau(y) \quad \text{for } (x,y) \in \cup\sigma\times\cup\tau,$$

where $M = m^d$ and $\mathcal{L}_{m_1}^\sigma$ and $\mathcal{L}_{m_2}^\tau$ are the appropriate tensor Lagrange polynomials, this implies

$$\widetilde{A}_{jk}^{\alpha\beta} = \int_{T_j}\int_{T_k} \widetilde{\kappa}_{\alpha\beta}(x,y)\,dy\,dx = \sum_{m_1=0}^{M}\sum_{m_2=0}^{M} \underbrace{\kappa_{\alpha\beta}\big(x_{m_1}^\sigma, x_{m_2}^\tau\big)}_{=:S_{m_1m_2}^{\alpha\beta\sigma\tau}} \times \underbrace{\int_{T_j}\mathcal{L}_{m_1}^\sigma(x)\,dx}_{=:V_{jm_1}^\sigma} \underbrace{\int_{T_k}\mathcal{L}_{m_2}^\tau(y)\,dy}_{=:V_{km_2}^\tau}$$

for $T_j \in \sigma$ and $T_k \in \tau$. With the matrices $V^\sigma \in \mathbb{R}^{|\sigma|\times M}$, $V^\tau \in \mathbb{R}^{|\tau|\times M}$, and $S^{\alpha\beta\sigma\tau} \in \mathbb{R}^{M\times M}$, the submatrix $\mathbf{A}^{\alpha\beta}|_{\sigma\times\tau}$ of $\mathbf{A}^{\alpha\beta}$ can be computed approximately by a matrix product

$$\mathbf{A}^{\alpha\beta}|_{\sigma\times\tau} \approx V^\sigma S^{\alpha\beta\sigma\tau} V^{\tau\,T} =: \widetilde{\mathbf{A}}^{\alpha\beta}|_{\sigma\times\tau}. \tag{57}$$

Second, for an inadmissible block $(\sigma,\tau) \in \mathbb{P}_{\mathrm{near}}$, we have $\widetilde{\mathbf{A}}^{\alpha\beta}|_{\sigma\times\tau} = a(\chi_{T_j}\mathbf{e}_\alpha, \chi_{T_k}\mathbf{e}_\beta) = \mathbf{A}^{\alpha\beta}|_{\sigma\times\tau}$; these entries are computed by (47). Double boundary integrals as in (47) occur in the context of boundary integral methods with piecewise constant ansatz and test functions. For simple geometries of the elements, analytic formulae are known, cf. [Hac02, Mai99, Mai00].

***Remark* 10.** Note that only the multiplication matrices $S^{\alpha\beta\sigma\tau}$ and $\mathbf{A}^{\alpha\beta}$ in (57) do depend on the indices α,β. Therefore, on admissible blocks $(\sigma,\tau) \in \mathbb{P}_{\mathrm{far}}$, the matrices $\widetilde{\mathbf{A}}^{\alpha\beta}|_{\sigma\times\tau}$ should be treated simultaneously for all $1 \le \alpha \le \beta \le d$.

***Remark* 11.** For inadmissible blocks $(\sigma,\tau) \in \mathbb{P}_{\mathrm{near}}$, the matrices $\widetilde{\mathbf{A}}^{\alpha\beta}|_{\sigma\times\tau} = \mathbf{A}^{\alpha\beta}|_{\sigma\times\tau}$ should also be assembled simultaneously, as the entries only differ on the components of the normal vectors, cf. (47). Since the block partitioning is symmetric and we are using constant approximation order, $\widetilde{\mathbf{A}}^{\alpha\beta}$ also is symmetric; therefore, $\widetilde{A}_{jk}^{\alpha\beta}$ should only be computed and stored for $1 \le j \le k \le N$.

5.2. Approximation of A on Far Field Blocks via (47). A different $\mathcal{H}^2$-matrix approach can be realized by use of (47) and by replacing G by $\mathcal{I}_m^{(\sigma,\tau)}G$ on admissible blocks. We start out from (54) and consider the following discretization of the bilinear form $a(\cdot,\cdot)$, where $\kappa(x,y) := G(x-y)$ now:

$$\begin{aligned}\widetilde{a}(\mathbf{m}_h,\widetilde{\mathbf{m}}_h)&\\ &= -\sum_{(\sigma,\tau)\in\mathbb{P}_{\text{near}}}\sum_{T_j\in\sigma}\sum_{T_k\in\tau}\int_{\partial T_j}\int_{\partial T_k}\big(\boldsymbol{\nu}^j(x)\cdot\mathbf{m}_h^j\big)\kappa(x,y)\big(\boldsymbol{\nu}^k(y)\cdot\widetilde{\mathbf{m}}_h^k\big)\,ds_y\,ds_x\\ &\quad-\sum_{(\sigma,\tau)\in\mathbb{P}_{\text{far}}}\sum_{T_j\in\sigma}\sum_{T_k\in\tau}\int_{\partial T_j}\int_{\partial T_k}\big(\boldsymbol{\nu}^j(x)\cdot\mathbf{m}_h^j\big)\mathcal{I}_m^{(\sigma,\tau)}\kappa(x,y)\big(\boldsymbol{\nu}^k(y)\cdot\widetilde{\mathbf{m}}_h^k\big)\,ds_y\,ds_x.\end{aligned}$$

We state the implementational details first. The only difference to Section **5.1** is in the way how $\widetilde{\mathbf{A}}$ is computed on admissible blocks $(\sigma,\tau)\in\mathbb{P}_{\text{far}}$. Given the degenerate kernel

$$\mathcal{I}_m^{(\sigma,\tau)}\kappa(x,y) = \sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\kappa\big(x_{m_1}^\sigma,x_{m_2}^\tau\big)\mathcal{L}_{m_1}^\sigma(x)\mathcal{L}_{m_2}^\tau(y),$$

we obtain for $T_j\in\sigma$, $T_k\in\tau$ using integration by parts

$$\begin{aligned}\widetilde{A}_{jk}^{\alpha\beta} = \widetilde{a}(\chi_{T_j}\mathbf{e}_\alpha,\chi_{T_k}\mathbf{e}_\beta) &= -\int_{\partial T_j}\int_{\partial T_k}\boldsymbol{\nu}_\alpha^j(x)\mathcal{I}_m^{(\sigma,\tau)}\kappa(x,y)\boldsymbol{\nu}_\beta^k(y)\,ds_y\,ds_x\\ &= -\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\kappa\big(x_{m_1}^\sigma,x_{m_2}^\tau\big)\int_{\partial T_j}\mathcal{L}_{m_1}^\sigma(x)\boldsymbol{\nu}_\alpha^j(x)\,ds_x\int_{\partial T_k}\mathcal{L}_{m_2}^\tau(y)\boldsymbol{\nu}_\beta^k(y)\,ds_y\\ &= -\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\underbrace{\kappa\big(x_{m_1}^\sigma,x_{m_2}^\tau\big)}_{=S_{m_1m_2}^{\sigma\tau}}\underbrace{\int_{T_j}\frac{\partial\mathcal{L}_{m_1}^\sigma}{\partial x_\alpha}\,dx}_{=:V_{jm_1}^{\sigma\alpha}}\underbrace{\int_{T_k}\frac{\partial\mathcal{L}_{m_2}^\tau}{\partial y_\beta}\,dy}_{=:V_{km_2}^{\tau\beta}}.\end{aligned}$$

***Remark* 12.** The integrands in $V_{jm_1}^{\sigma\alpha}$ and $V_{km_2}^{\tau\beta}$ are computable, as formulae for the first derivatives of the Lagrange polynomials in one dimension can easily be derived by induction: given $\mathcal{L}_j\in\mathcal{P}_m$, we have

$$\mathcal{L}_j'(t) = \sum_{\substack{k=1\\k\neq j}}^{m}\frac{1}{t_j-t_k}\prod_{\substack{\ell=1\\\ell\notin\{j,k\}}}^{m}\frac{t-t_\ell}{t_j-t_\ell}.$$

***Remark* 13.** Note that the multiplication matrices $S^{\sigma\tau}$ do *not* depend on α,β here. Thus, we now have to assemble and store $S^{\sigma\tau}$, the matrices $V^{\sigma\alpha}$ and $V^{\tau\beta}$ for all leaves σ,τ of the cluster tree only, and the father-son transformation matrices $B^{\tau'\tau}$. This is a significant difference to Section **5.1**, where the computation of $S^{\alpha\beta\sigma\tau}$, of V^τ for all leaves τ, and of $B^{\tau'\tau}$ is required.

The results of Section **5.1** now carry over immediately; however, we have to assume that m is large enough so that

$$\|\partial_\alpha\partial_\beta(\kappa-\mathcal{I}_m^{(\sigma,\tau)}\kappa)\|_{L^\infty(B_\sigma\times B_\tau)}\le c_3$$

for $\kappa(x,y)=G(x-y)$, with a constant c_3 which is independent of $(\sigma,\tau)\in\mathbb{P}_{\text{far}}$ and $1\le\alpha,\beta\le d$. The existence of such a constant is a consequence of Lemma 3.6. In complete analogy to Theorem 5.2 we conclude:

Theorem 5.3. *Under the above assumptions, we have*

$$|a(\mathbf{m}_h, \widetilde{\mathbf{m}}_h) - \widetilde{a}(\mathbf{m}_h, \widetilde{\mathbf{m}}_h)| \leq c_3 d|\Omega| \|\mathbf{m}_h\|_{L^2(\Omega)} \|\widetilde{\mathbf{m}}_h\|_{L^2(\Omega)} \tag{58}$$

for all $\mathbf{m}_h, \widetilde{\mathbf{m}}_h \in \mathcal{P}_0(\mathcal{T})^d$. □

5.3. Solvability of the Approximate Discrete Model $(\widetilde{RP}_{\varepsilon,h})$**.** Suppose a triangulation $\mathcal{T}$ of Ω by rectangular, axis-parallel boxes to be given. The Gauss Divergence Theorem then shows that the restriction of $\mathcal{P}$, $\mathcal{P}|_{\mathcal{P}_0(\mathcal{T})^d}$, is injective, cf. [CP04a]. In particular, there exists a constant c_4 such that

$$a(\mathbf{m}_h, \mathbf{m}_h) = \|\mathcal{P}\mathbf{m}_h\|^2_{L^2(\mathbb{R}^d)} \geq c_4 \|\mathbf{m}_h\|^2_{L^2(\Omega)} \quad \text{for all } \mathbf{m}_h \in \mathcal{P}_0(\mathcal{T})^d.$$

Theorem 5.4. *Given the above assumptions, take m large enough so that $C_{\rm df} := c_4 - c_3 d|\Omega| \geq 0$, with c_3 from Theorems 5.2 and 5.3, respectively. Then, the approximate discrete model $(\widetilde{RP}_{\varepsilon,h})$ has solutions. In case $C_{\rm df} > 0$, the solution to $(\widetilde{RP}_{\varepsilon,h})$ is unique.*

Proof. As

$$\widetilde{a}(\mathbf{m}_h, \mathbf{m}_h) \geq a(\mathbf{m}_h, \mathbf{m}_h) - |a(\mathbf{m}_h, \mathbf{m}_h) - \widetilde{a}(\mathbf{m}_h, \mathbf{m}_h)| \geq C_{\rm df} \|\mathbf{m}_h\|^2_{L^2(\Omega)},$$

one concludes with $C_{\rm df} \geq 0$ that the approximate bilinear form $\widetilde{a}(\cdot,\cdot)$ is positive semidefinite, and even positive definite if $C_{\rm df} > 0$. Thus, the symmetry of $\widetilde{a}(\cdot,\cdot)$ shows that $\mathbf{m}_h \mapsto \widetilde{a}(\mathbf{m}_h, \mathbf{m}_h)$ is a convex functional. In sum, the approximate energy functional

$$\widetilde{E}(\mathbf{m}_h) := \int_\Omega \phi^{**}(\mathbf{m}_h)\, dx - \int_\Omega \mathbf{f} \cdot \mathbf{m}_h \, dx + \frac{1}{2}\widetilde{a}(\mathbf{m}_h, \mathbf{m}_h) + \frac{1}{2}\int_\Omega \varepsilon^{-1}(|\mathbf{m}_h| - 1)^2_+ \, dx \tag{59}$$

is continuous and convex. Coercivity of $\widetilde{E}$ follows from the last term in (59), i.e., from the penalization energy contribution.

The Direct Method of the Calculus of Variations now proves the existence of (global) minimizers for $\widetilde{E}(\cdot)$. As minimizers of $\widetilde{E}$ in $\mathcal{P}_0(\mathcal{T})^d$ are zeros of the corresponding Gâteaux derivatives, which read $(\widetilde{RP}_{\varepsilon,h})$, we have thus shown the solvability of $(\widetilde{RP}_{\varepsilon,h})$.

To prove uniqueness in case $C_{\rm df} > 0$, assume $(\lambda_h, \mathbf{m}_h)$ and $(\widetilde{\lambda}_h, \widetilde{\mathbf{m}}_h)$ to be two solutions of $(\widetilde{RP}_{\varepsilon,h})$ now. One immediately finds

$$\begin{aligned} \widetilde{a}(\mathbf{m}_h - \widetilde{\mathbf{m}}_h, \mathbf{m}_h - \widetilde{\mathbf{m}}_h) + \langle D\phi^{**}(\mathbf{m}_h) - D\phi^{**}(\widetilde{\mathbf{m}}_h) \,;\, \mathbf{m}_h - \widetilde{\mathbf{m}}_h \rangle_{L^2(\Omega)} \\ + \langle \lambda_h \mathbf{m}_h - \widetilde{\lambda}_h \widetilde{\mathbf{m}}_h \,;\, \mathbf{m}_h - \widetilde{\mathbf{m}}_h \rangle_{L^2(\Omega)} = 0. \end{aligned} \tag{60}$$

As ϕ^{**} is convex, the second term in (60) cannot be negative, whereas the third term is non-negative by a direct calculation [CP01]. Hence, $\widetilde{a}(\mathbf{m}_h - \widetilde{\mathbf{m}}_h, \mathbf{m}_h - \widetilde{\mathbf{m}}_h) = 0$, which implies $\mathbf{m}_h = \widetilde{\mathbf{m}}_h$ due to the definiteness of $\widetilde{a}(\cdot,\cdot)$. □

***Remark* 14.** The a priori and a posteriori error analysis for the approximate model $(\widetilde{RP}_{\varepsilon,h})$ is the topic of ongoing research and will appear in a subsequent work.

6. Numerical Experiments

In the following section we collect the results of our numerical experiments. We compare the performance of the two $\mathcal{H}^2$-matrix approximations introduced in Sections **5.1** and **5.2** to that of the standard approach using the full stiffness matrices.

6.1. **Implementational Details.** All numerical experiments were conducted using the HLib software package provided by S. Börm and L. Grasedyck of the Max-Planck-Institute for Mathematics in the Sciences (Leipzig). We utilized a Compaq/HP AlphaServer ES45 under Unix, with four Alpha EV68 CPUs running at 1 GHz each and 32 GBytes of RAM.

For our experiments, we varied the interpolation order m in the $\mathcal{H}^2$-matrix approximation between 2 and 6. As for the parameters, we fixed the maximum leaf size $C_{\ell\mathrm{f}} = 20$ and the admissibility parameter $\eta = 1$ throughout. However, to avoid having to vary $C_{\ell\mathrm{f}}$ with m, we decided to store admissible blocks $(\sigma, \tau) \in \mathbb{P}_{\mathrm{far}}$ as full matrices whenever $|\sigma||\tau| \leq M_\sigma M_\tau$, i.e., whenever storing the full matrix was less expensive than storing the corresponding multiplication matrix $S^{\sigma\tau}$.
Moreover, to be able to use the supplied HLib routines with as few modifications as possible, we did exploit the symmetry of $\mathbf{A}^{\alpha\beta}$ when setting up the matrices, but neglected it in the process of storing them.

6.2. **Full Matrices vs. $\mathcal{H}^2$-Matrices.** In a first example, we restrict ourselves to a comparison of the properties of $\mathbf{A}$ and its $\mathcal{H}^2$-approximation $\widetilde{\mathbf{A}}$. For simplicity, we assume the domain Ω to be the unit square, $\Omega := [0,1] \times [0,1] \subset \mathbb{R}^2$, and consider a uniform triangulation of Ω consisting of rectangular elements. The number N of degrees of freedom is varied between 256 and 1048576. Note that for $N \geq 16384$, we have no longer set up the full matrix $\mathbf{A}$, but have approximated it by an $\mathcal{H}^2$-matrix with $m = 10$; all subsequent references are made to this approximation.

We compare the results of our two $\mathcal{H}^2$-matrix approaches for fixed approximation orders m.
First, we give the relative approximation errors $\|\mathbf{A} - \widetilde{\mathbf{A}}\|_2/\|\mathbf{A}\|_2$; the values are collected in the following Tables 1 and 2.

TABLE 1. Relative approximation errors (first $\mathcal{H}^2$-matrix approach, cf. Section **5.1**).

N/m	2	3	4	5	6
256	2.38605_{-3}	2.42547_{-4}	2.22014_{-16}	2.36817_{-16}	2.10329_{-16}
1024	4.66221_{-3}	4.74404_{-4}	4.42905_{-5}	4.45515_{-6}	3.85638_{-7}
4096	6.08654_{-3}	6.08069_{-4}	6.14339_{-5}	5.88869_{-6}	5.21420_{-7}
16384	7.28966_{-3}	7.41998_{-4}	7.76218_{-5}	7.11627_{-6}	7.26753_{-7}
65536	8.42299_{-3}	8.57423_{-4}	8.83683_{-5}	8.26648_{-6}	8.21227_{-7}

TABLE 2. Relative approximation errors (second $\mathcal{H}^2$-matrix approach, cf. Section **5.2**).

N/m	2	3	4	5	6
256	1.47377_{-2}	1.80771_{-3}	2.46549_{-16}	2.25802_{-16}	2.15506_{-16}
1024	3.80124_{-2}	3.47728_{-3}	2.38627_{-4}	2.65527_{-5}	2.05902_{-6}
4096	5.48085_{-2}	4.98563_{-3}	4.17924_{-4}	3.81013_{-5}	3.65331_{-6}
16384	6.54358_{-2}	6.24201_{-3}	5.19552_{-4}	5.21367_{-5}	5.78795_{-6}
65536	7.10886_{-2}	7.69041_{-3}	6.14949_{-4}	6.02877_{-5}	6.88747_{-6}

In summary, the errors in the second approach (cf. Section **5.2**) seem to be larger by one order of magnitude. The convergence rates, however, are optimal in both cases: every increase of m by one reduces the error by an order of magnitude. Note that for $N = 256$ and $m \geq 4$, our choice of $C_{\ell\mathrm{f}}$ implies that there are no admissible blocks; the error in this case is due to rounding. The error estimates themselves are computed by a power iteration, with a maximum of 100 iterative steps.

Second, we consider the times needed for building the two $\mathcal{H}^2$-matrix approximations and compare them to the setup times of the full matrix $\mathbf{A}$, see Tables 3 and 4.

TABLE 3. Setup time in seconds (first approach).

N/m	2	3	4	5	6	full
256	0.7	0.7	0.8	0.8	0.8	0.8
1024	3.8	3.8	9.0	9.0	10.3	13.5
4096	18.1	18.0	50.9	51.1	60.1	215.6
16384	78.2	78.5	237.4	237.9	286.1	n/a
65536	329.0	330.9	1037.3	1042.2	1263.9	n/a
262144	1353.3	1361.8	4303.4	4308.8	5248.6	n/a
1048576	5421.2	5925.1	n/a	n/a	n/a	n/a

TABLE 4. Setup time in seconds (second approach).

N/m	2	3	4	5	6	full
256	0.7	0.6	0.8	0.9	0.9	0.8
1024	3.8	3.9	9.1	9.1	10.4	13.5
4096	18.1	18.0	51.1	51.4	60.7	215.6
16384	78.6	78.4	238.0	238.5	287.7	n/a
65536	330.1	331.5	1036.8	1037.3	1262.8	n/a
262144	1343.5	1349.4	4291.0	4311.0	5259.6	n/a
1048576	5450.3	5445.4	n/a	n/a	n/a	n/a

The time required for setting up the full matrix by far surpasses the setup times of the approximations; the gap increases with the number N of degrees of freedom. For m fixed, the difference between the two approaches is negligible here.

Third, of particular interest is the amount of memory required for storing the matrix approximations, as compared to the storage requirements of the original matrices; these are listed in Tables 5 and 6.

Overall, the figures compare very favourably with the storage required by the full matrices; here, the second approach is clearly superior to the first one for m fixed and N large. This is probably due to the fact that the multiplication matrices on admissible blocks only have to be stored once instead of for each block. Note that on meshes with N small, the memory requirements slightly favour the full matrix approach. This is due to the organizational effort involved in constructing the cluster tree of the $\mathcal{H}^2$-matrix approximation and in allocating memory blockwise for subblocks instead of for the whole matrix.

TABLE 5. Memory requirement in KBytes/N (first approach).

N/m	2	3	4	5	6	full
256	5.2	5.6	6.8	7.2	7.6	6.0
1024	8.8	11.6	18.8	20.5	22.1	24.0
4096	11.4	16.2	28.9	34.6	42.6	96.0
16384	12.9	19.1	35.5	44.6	58.2	384.0
65536	13.7	20.7	39.3	50.6	67.9	1536.0
262144	14.2	21.6	41.3	53.9	73.4	6144.0
1048576	14.4	22.1	n/a	n/a	n/a	24576.0

TABLE 6. Memory requirement in KBytes/N (second approach).

N/m	2	3	4	5	6	full
256	5.2	5.5	7.2	8.1	9.2	6.0
1024	8.3	9.4	18.6	19.9	22.5	24.0
4096	10.4	12.2	26.9	29.5	35.4	96.0
16384	11.6	13.8	31.8	35.6	44.1	384.0
65536	12.2	14.8	34.6	39.1	49.2	1536.0
262144	12.6	15.2	36.0	40.9	51.9	6144.0
1048576	12.7	15.5	n/a	n/a	n/a	24576.0

Altogether, these numerical experiments underline the applicability of $\mathcal{H}$-matrix techniques to the discretized potential operator $\mathcal{P}$ from $(RP_{\varepsilon,h})$. For given approximation order m, the first approach leads to lesser approximation errors, but is otherwise also more costly numerically, as is reflected by the much higher storage requirements.

In a certain sense, however, the two approaches seem almost equivalent: if we require some fixed accuracy, the approximation order m always has to be higher by one in the second approach, as the errors lag behind by one order of magnitude. A comparison of the respective setup times and memory requirements then shows the numerical cost to be almost even.

6.3. An Example with Known Exact Solution. In our second example, we consider a model problem for the relaxed Landau-Lifshitz problem (RP) taken from [CP04b]. As above, let the domain Ω be the unit square; assume Ω to be filled with some uniaxial magnetized material, with the easy axis given by $\mathbf{e} = (-1,1)/\sqrt{2}$ and the corresponding normal by $\mathbf{z} = (1,1)/\sqrt{2}$, see Remark 1. Define $(\mathbf{m}, \lambda) \in W^{1,\infty}(\Omega;\mathbb{R}^2) \times L^\infty(\Omega)$ as

$$\mathbf{m}(x) := \begin{cases} x & \text{for } |x| \le 1, \\ x/|x| & \text{for } |x| > 1 \end{cases} \quad \text{and} \quad \lambda(x) := \begin{cases} 0 & \text{for } |x| \le 1, \\ 1 & \text{for } |x| > 1. \end{cases} \tag{61}$$

Then, $(\mathbf{m}, \lambda)$ solves (RP) with given right-hand side

$$\mathbf{f} := \mathcal{P}\mathbf{m} + (\mathbf{m}\cdot\mathbf{z})\mathbf{z} + \lambda\mathbf{m} \quad \text{in } L^2(\Omega;\mathbb{R}^2), \tag{62}$$

cf. (11),(12). In the following, we replace $\mathcal{P}\mathbf{m}$ in (62) by $\mathcal{P}\mathbf{m}_{\mathcal{T}}$, where $\mathbf{m}_{\mathcal{T}}$ denotes the piecewise integral mean of $\mathbf{m}$. Note that there are no fully analytic examples

for (RP) with known solutions, which is why we have to restrict ourselves to the present model.

As $\mathbf{m}$ is Lipschitz continuous and therefore in $W^{1,\infty}(\Omega;\mathbb{R}^2)$, the a priori analysis from [CP04b] yields $\|(\mathbf{m}-\mathbf{m}_h)\cdot\mathbf{z}\|_{L^2(\Omega)} = \mathcal{O}(\varepsilon+h)$, with ε the penalization parameter from (14). For our experiments, we choose $\varepsilon = h$ and compute the discrete solution $\mathbf{m}_h = \sum_{j=1}^{2N} x_j\varphi_j$ with respect to the basis $\{\varphi_1,\dots,\varphi_{2N}\}$ from (50) by a classical Newton-Raphson scheme: the unknown coefficient vector $\mathbf{x}\in\mathbb{R}^{2N}$ is determined as the unique zero of

$$F(\mathbf{x}) := \big(\langle \mathcal{P}\mathbf{m}_h + D\phi^{**}(\mathbf{m}_h) + \lambda_h\mathbf{m}_h - \mathbf{f}\,;\,\varphi_j\rangle\big)_{j=1}^{2N} = 0. \tag{63}$$

[A detailed discussion on the relevance of ε can be found in [CP04b].] Note that the convergence of the Newton-Raphson method is not guaranteed mathematically, since F is only differentiable almost everywhere. The Jacobian of F can be written as a finite sum $DF(\mathbf{x}) = \mathbf{A} + \sum_{j=1}^{N} D_j(\mathbf{x})$ with symmetric positive semidefinite matrices $D_j(\mathbf{x})\in\mathbb{R}^{2N\times 2N}$, cf. [CP04a]. For a triangulation by rectangular elements such as ours, the matrix $\mathbf{A}$ can be shown to be positive definite by applying the Gauss Divergence Theorem, see [CP04b]. We therefore employ a preconditioned conjugate gradient method, with the LU decomposition of a coarsened $\mathcal{H}$-matrix version of $\widetilde{\mathbf{A}}$ as preconditioner.

TABLE 7. Number of Newton steps: first approach (left), second approach (right).

N/m	2	3	4	5	6
256	6	6	6	6	6
1024	8	8	8	8	8
4096	8	8	8	8	8
16384	9	9	9	9	9

N/m	2	3	4	5	6
256	6	6	6	6	6
1024	8	8	8	8	8
4096	7	8	8	8	8
16384	8	8	9	9	9

In Table 7, we summarize the number of Newton steps required for finding the coefficient vector $\mathbf{x}$ of $\mathbf{m}_h$ in (63). One sees that the number of steps is nearly constant, i.e., independent of m and growing only slightly with N.

Figure 2 gives the convergence history of the full error $\|\mathbf{m}-\mathbf{m}_h\|_{L^2(\Omega)}$ for different choices of the order m in the $\mathcal{H}^2$-matrix approximation. The convergence rate is almost the optimal $1/2$, up to minimal deviations which are presumably due to the fact that we approximate $\mathcal{P}\mathbf{m}$ by $\mathcal{P}\mathbf{m}_{\mathcal{T}}$ in (62). As was to be expected, the interpolation order m has to increase with the number N of degrees of freedom in order to maintain optimal convergence; the effect is clearly more pronounced for the second approach.

Note that we consider the full L^2 norm instead of $\|D\phi^{**}(\mathbf{m}) - D\phi^{**}(\mathbf{m}_h)\|_{L^2(\Omega)}$, although only the latter is covered by the available a priori error analysis. Recently it has been shown that one always obtains weak L^2 convergence of $\mathbf{m}_h \rightharpoonup \mathbf{m}$; more precisely, there holds $\|\mathbf{m}-\mathbf{m}_h\|_{\widetilde{H}^{-1}(\Omega)} = \mathcal{O}(\varepsilon+h)$, cf. [CP04c] . Nevertheless, the numerical experiments available from [CP04b] indicate that one may hope for sharper results.

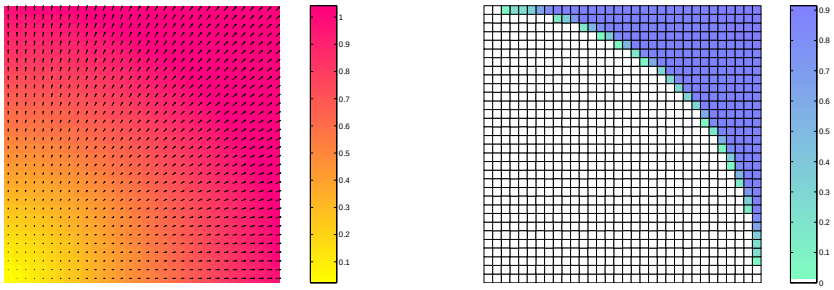

FIGURE 1. The discrete solution $(\mathbf{m}_h, \lambda_h)$ of $(RP_{\varepsilon,h})$ as in the model example for $N = 1024$, with $\mathbf{m}_h$ on the left (displayed as vectors $\mathbf{m}_h|_T$ and $|\mathbf{m}_h|_T|$) and λ_h on the right. In the white region, we have $|\mathbf{m}_h| \leq 1$ and therefore $\lambda_h = 0$.

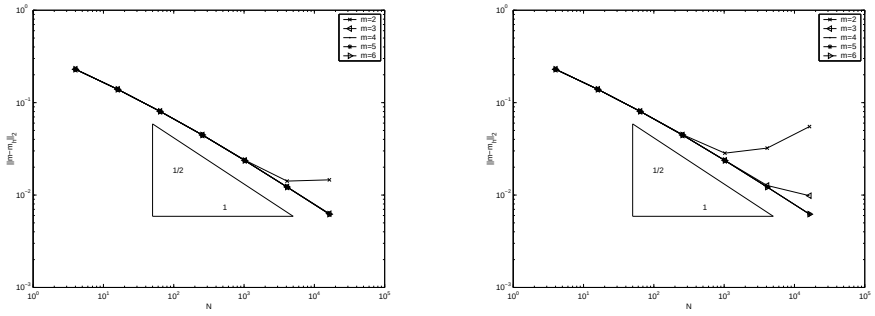

FIGURE 2. Experimental convergence of $\|\mathbf{m} - \mathbf{m}_h\|_{L^2(\Omega)}$ over N: first approach (left), second approach (right).

Acknowledgement. This work has been supported by the Austrian Science Fund FWF under grant P15274, "*Reliable and Efficient Numerical Algorithms in Computational Micromagnetics*". The authors thank S. Börm and L. Grasedyck of the Max-Planck-Institute for Mathematics in the Sciences (Leipzig) for providing them with the HLib software package, as well as for several helpful comments.

REFERENCES

[BG04] S. Börm and L. Grasedyck. Low-Rank Approximation of Integral Operators by Interpolation. *Computing*, 72:325–332, 2004.

[BGH03] S. Börm, L. Grasedyck, and W. Hackbusch. Introduction to Hierarchical Matrices with Applications. *Eng. Anal. with B. Elem.*, 27:405–422, 2003.

[BH02] S. Börm and W. Hackbusch. $\mathcal{H}^2$-Matrix Approximation of Integral Operators by Interpolation. *Appl. Numer. Math.*, 43:129–143, 2002.

[BLM04] S. Börm, M. Löhndorf, and J.M. Melenk. Approximation of Integral Operators by Variable-Order Interpolation. *Numer. Math.*, 2004. To appear.

[BS01] I. Babuška and T. Strouboulis. *The Finite Element Method and its Reliability*. Numerical Mathematics and Scientific Computation. Oxford University Press, Inc., Oxford, 2001.

[CP01] C. Carstensen and A. Prohl. Numerical analysis of relaxed micromagnetics by penalised finite elements. *Numer. Math.*, 90:65–99, 2001.

[CP04a] C. Carstensen and D. Praetorius. Effective Simulation of a Macroscopic Model for Stationary Micromagnetics. *Comput. Methods Appl. Mech. Engrg.*, 2004. To appear.

[CP04b] C. Carstensen and D. Praetorius. Numerical Analysis for a Macroscopic Model in Micromagnetics. *SIAM J. Numer. Anal.*, 2004. To appear.

[CP04c] C. Carstensen and D. Praetorius. Stabilization Yields Strong Convergence for a Macroscopic Model in Micromagnetics. 2004. In preparation.

[DeS93] A. DeSimone. Energy Minimizers for Large Ferromagnetic Bodies. *Arch. Rational Mech. Anal.*, 125:99–143, 1993.

[DL93] R.A. DeVore and G.G. Lorentz. *Constructive Approximation*, volume 303 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin Heidelberg, 1993.

[GH03] L. Grasedyck and W. Hackbusch. Construction and arithmetics of $\mathcal{H}$-matrices. *Computing*, 70(4):295–334, 2003.

[Gie01] K. Giebermann. Multilevel Approximation of Boundary Integral Operators. *Computing*, 67:183–207, 2001.

[Gra01] L. Grasedyck. *Theorie und Anwendungen Hierarchischer Matrizen*. PhD thesis, Christian-Albrechts-Universität, Kiel, 2001.

[Hac02] W. Hackbusch. Direct Integration of the Newton Potential over Cubes Including a Program Description. *Computing*, 68:193–216, 2002.

[HKS00] W. Hackbusch, B. Khoromskij, and S.A. Sauter. On $\mathcal{H}^2$-matrices. In *Lectures on Applied Mathematics*, pages 9–29. Springer-Verlag, Berlin Heidelberg, 2000.

[HS98] A. Hubert and R. Schäfer. *Magnetic Domains: the Analysis of Magnetic Microstructures*. Springer-Verlag, Berlin Heidelberg, 1998.

[JK90] R.D. James and D. Kinderlehrer. Frustration in ferromagnetic materials. *Cont. Mech. Therm.*, 2:215–239, 1990.

[KP01] M. Kružík and A. Prohl. Young measure approximation in micromagnetics. *Numer. Math.*, 90(2):291–307, 2001.

[LM92] M. Luskin and L. Ma. Analysis of the finite element approximation of microstructure in micromagnetics. *SIAM J. Numer. Anal.*, 29(2):320–331, 1992.

[Ma91] L. Ma. *Analysis and Computation for a Variational Problem in Micromagnetics*. PhD thesis, University of Minnesota, Minneapolis, 1991.

[Mai99] M. Maischak. The Analytical Computation of the Galerkin Elements for the Laplace, Lamé and Helmholtz Equation in 2D-BEM. Technical Report ifam48, Institut für Angewandte Mathematik, Universität Hannover, 1999.

[Mai00] M. Maischak. The Analytical Computation of the Galerkin Elements for the Laplace, Lamé and Helmholtz Equation in 3D-BEM. Technical Report ifam50, Institut für Angewandte Mathematik, Universität Hannover, 2000.

[Ped97] P. Pedregal. *Parametrized Measures and Variational Principles*. Birkhäuser, Basel, 1997.

[Pra04] D. Praetorius. Analysis of the Operator $\triangle^{-1}$div Arising in Magnetic Models. *ZAA*, 23:589–605, 2004.

[Pro01] A. Prohl. *Computational micromagnetism*. Advances in Numerical Mathematics. B.G. Teubner, Stuttgart, 2001.

[Riv84] T.J. Rivlin. *The Chebyshev Polynomials*. Wiley Interscience, New York, 1984.

[Tar95] L. Tartar. Beyond Young Measures. *Meccanica*, 30:505–526, 1995.

E-mail address: Nikola.Popovic@tuwien.ac.at, Dirk.Praetorius@tuwien.ac.at

Vienna University of Technology, Institute for Analysis and Scientific Computing, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria