Numerical quadratic energy minimization bound to convex constraints in thin-film micromagnetics

S. Ferraz-Leite, JM. Melenk, D. Praetorius

Most recent ASC Reports

Numerical quadratic energy minimization bound to convex constraints in thin-film micromagnetics

Samuel Ferraz-Leite[1], Jens Markus Melenk[2], and Dirk Praetorius[2]

[1] Max-Planck-Institute for Mathematics in the Sciences
Inselstraße 22, 04103 Leipzig, Germany

samuel.ferraz-leite@mis.mpg.de (corresponding author)

[2] Vienna University of Technology
Wiedner Hauptstraße 8-10, A-1040 Vienna, Austria

{melenk,dirk.praetorius}@tuwien.ac.at

August 29, 2011

Abstract

We analyze the reduced model for thin-film devices in stationary micromagnetics proposed in [7, DeSimone, Kohn, Müller, Otto, Schäfer 2001]. We introduce an appropriate functional analytic framework and prove well-posedness of the model in that setting. The scheme for the numerical approximation of solutions consists of two ingredients: The energy space is discretized in a conforming way using Raviart-Thomas finite elements; the non-linear but convex side constraint is treated with a penalty method. This strategy yields a convergent sequence of approximations as discretization and penalty parameter vanish. The proof generalizes to a large class of minimization problems and is of interest beyond the scope of thin-film micromagnetics. Numerical experiments support our findings and illustrate the performance of the proposed algorithm.

Keywords: quadratic programming, convex constraints, penalty method, thin-film micromagnetics
Mathematics Subject Classification (2010): 65K05, 65K15, 49M20

1 Introduction and Abstract Setting

Let $\Omega \subseteq \mathbb{R}^n$ be a domain and $\mathcal{H} \subseteq L^2(\Omega)$ denote a continuously embedded Hilbert space with norm $\|u\|_{L^2(\Omega)}^2 \lesssim \|u\|_{\mathcal{H}}^2 = (u, u)_{\mathcal{H}}$. Let $\|u\|^2 = \langle\!\langle u, u \rangle\!\rangle$ be a continuous semi-norm on the space $\mathcal{H}$ which is induced by the bilinear form $\langle\!\langle \cdot, \cdot \rangle\!\rangle$. Let $g : \mathcal{H} \to L^2(\Omega)$ be a continuous mapping that is defined on the whole space $\mathcal{H}$. We assume that g is pointwise convex, i.e. for two functions $u, v \in \mathcal{H}$ and for any choice of $\lambda \in (0, 1)$, the inequality

$$g(\lambda u + (1 - \lambda)v) \leq \lambda g(u) + (1 - \lambda)g(v)$$

holds true pointwise almost everywhere. Given some linear and continuous functional $\Phi : \mathcal{H} \to \mathbb{R}$, we define the quadratic energy functional

$$e(u) = \frac{1}{2}\|u\|^2 - \Phi(u). \tag{1}$$

Minimization problem (M): Find a minimizer u^* of $e(\cdot)$ subject to the convex side constraint

$$g(u^*) \leq 0 \in L^2(\Omega). \tag{2}$$

The side constraint (2) defines a closed and convex set $\mathcal{A} := \{u \in \mathcal{H} \mid g(u) \leq 0 \text{ a.e.}\}$ of admissible functions. As a general assumption, we state the coercivity of the energy functional on the set of

admissible functions:

$$\lim_{\substack{\|u\|_{\mathcal{H}} \to \infty \\ u \in \mathcal{A}}} e(u) = \infty, \quad \inf_{u \in \mathcal{A}} e(u) > -\infty \tag{A1}$$

Under these assumptions the direct method of calculus of variations [5] provides existence of a minimizer $u^* \in \mathcal{A}$. Moreover, if $\|\cdot\|$ is a norm, the minimizer u^* is unique. The side constraint, however, may be difficult to treat numerically. Suitable optimality conditions—also referred to as KKT equations—may be hard to derive. A black box scheme for the numerical solution of this class of minimization problems is desirable, and we proceed to present a convergent penalty method.

We define the function

$$(u)_+(x) := \begin{cases} u(x), & \text{if } u(x) \geq 0 \\ 0, & \text{else} \end{cases}$$

pointwise almost everywhere. With a parameter $\varepsilon > 0$, we introduce the penalty energy

$$\frac{1}{2\varepsilon} \|(g(u))_+\|_{L^2(\Omega)}^2.$$

Minimizing the energy (1) over the set $\mathcal{A}$ of admissible functions can be interpreted as minimizing the energy functional

$$e_0(u) := \begin{cases} e(u), & \text{for } u \in \mathcal{A} \\ \infty, & \text{else} \end{cases}$$

over the full space $\mathcal{H}$. We observe that this functional is not smooth since it has a sharp energy barrier at the boundaries of $\mathcal{A}$. The idea of the penalty method is to approximate the energy functional e_0 by a regularization. Given $\varepsilon > 0$, we seek a minimizer $u^\varepsilon \in \mathcal{H}$ of the penalized energy

$$e_\varepsilon(u) := e(u) + \frac{1}{2\varepsilon} \|(g(u))_+\|_{L^2(\Omega)}^2.$$

Throughout, we assume coercivity of the penalized energy functional for each fixed $\varepsilon > 0$, which essentially is an additional assumption on the function g:

$$\lim_{\substack{\|u\|_{\mathcal{H}} \to \infty \\ u \in \mathcal{H}}} e_\varepsilon(u) = \infty, \quad \inf_{u \in \mathcal{H}} e_\varepsilon(u) > -\infty \qquad \forall \varepsilon > 0. \tag{A2}$$

In order to solve the unconstrained penalized minimization problem we need to discretize the energy space $\mathcal{H}$. Let $(h_n)_{n \in \mathbb{N}}$ be a positive zero sequence and let $(X_{h_n})_{n \in \mathbb{N}}$ be a sequence of finite dimensional subspaces of $\mathcal{H}$. We assume

$$X_{h_i} \subseteq X_{h_j} \text{ for } i \leq j \qquad \text{and} \qquad \overline{\bigcup_{n \in \mathbb{N}} X_{h_n}} = \mathcal{H} \tag{3}$$

These conditions state that the sequence of discrete spaces is nested and that for vanishing discretization parameter the discrete subspaces become dense in $\mathcal{H}$. Both of these assumptions are usually satisfied for regular discretizations based on some mesh $\mathcal{T}_{h_n}$ of the domain Ω, where $\mathcal{T}_{h_{n+1}}$ is a uniform refinement of $\mathcal{T}_{h_n}$.

Let u_h^0 denote a solution to the discretized constrained minimization problem. A solution to the penalized continuous problem is denoted by u_0^ε. Only u_h^ε, a solution to the discretized and penalized problem, is computed in practice. In Section 2, we prove that this numerical approach is justified in the sense that any choice of zero sequences $(\varepsilon_n, h_n) \to (0, 0)$ yields convergence of the computable quantities $u_n := u_{h_n}^{\varepsilon_n}$, i.e.

$$\lim_{n \to \infty} u_n = u^*$$

in an appropriate sense. To be precise, we prove convergence with respect to the weak topology of $\mathcal{H}$ as well as the topology induced by the semi-norm $\|\cdot\|$. Recall that the solution u^* is in general

2

not unique. We clarify our notion of convergence: Each subsequence of $(u_n)_{n \in \mathbb{N}}$ has a convergent subsequence $(u_{n_k})_{k \in \mathbb{N}}$ whose limit $u^* := \lim_k u_{n_k}$ is a minimizer. If u^* is uniquely determined, the full sequence converges $u_n \to u^*$.

We apply our numerical scheme to the minimization problem proposed in [7] for the simulation of thin-film devices in micromagnetics. For that model, we define a functional analytic framework and analyze existence and uniqueness of solutions as well as continuous dependence on the given data in Section 3. Since the model problem fits into our abstract setting, we may use the penalty method to treat the side constraint. In Section 4 we use Raviart-Thomas finite elements to discretize the energy space $\mathcal{H}$. The behavior of our algorithm is studied with numerical experiments.

There is a vast literature on penalty methods. None of the results we found, however, could be applied to the thin-film model in micromagnetics discussed in Section 3. For discrete minimization problems with finite dimensional energy space arising, for example, in the context of mathematical finance, penalty methods are well understood and convergence is established, see [21] and the references therein. One may therefore be tempted to first discretize the continuous problem and then apply a penalty method to the obtained discrete minimization problem. There is no obvious mathematical justification, however, that refinement of the discretization does not increase the penalty error introduced.

In the context of infinite dimensional problems, a lot of effort has been put into developing efficient numerical schemes for the solution of quadratic minimization problems bound to side constraints, see e.g. [15, 16] and the references therein. Since it is known that the system of equations of the penalized discrete problem becomes ill-conditioned as $\varepsilon \to 0$, it is natural to ask for other, more robust methods for specific applications. All of these methods, however, are based on properties of the corresponding KKT-system or on orthogonal projections onto the admissible set. Apparently the non-linearity of the function g and the non-local norm of our energy space are significant enhancements to the problem complexity, and we failed to transfer ideas from the literature to our model problem.

In some cases, as the large-body limit in micromagnetics or contact problems, penalty methods have been applied successfully. Even convergence in certain norms has been proven, see e.g. [2, 3, 19]. These convergence results, however, are based on KKT-like conditions and require knowledge of the nature of the Lagrange multipliers associated with the inequality constraints. Since the thin-film model problem under consideration resembles the one treated in [2, 3], as a first attempt we tried to transfer or modify the proofs in the publications mentioned to our setting. One crucial ingredient in the analysis of the works concerned with the large-body limit is, however, that the energy space is $L^2(\Omega)$. In particular its dual space consists of Lebesgue functions only. In our case the dual space includes distributions, which required us to develop a new approach to prove convergence of our method. The new approach applies to a large class of problems, and we are confident that it is of interest beyond the specific application in thin-film micromagnetics. It avoids use of KKT equations; it avoids estimates that require information about the Lagrange multipliers.

2 An abstract convergence result

Before stating the abstract convergence result, we introduce some notation and collect the necessary assumptions: Let $\varepsilon, h \geq 0$. We denote by (M) the continuous and constrained minimization problem with solution u^*. The discretized and penalized minimization problem is referred to as (M_h^ε) with solution u_h^ε. In case of $\varepsilon = 0$ we have a constrained—and possibly discrete—problem (M_h^0) with solution u_h^0; the case $h = 0$ corresponds to a continuous—and possibly penalized—problem (M_0^ε) with solution u_0^ε. Our notation identifies $(M_0^0) = (M)$ and $u_0^0 = u^*$. We stress that if $\|\cdot\|$ is a norm, all problems (M_h^ε) allow for a uniquely determined solution u_h^ε. In this case, the particularly interesting situation is the one where the energy norm $\|\cdot\|$ is *not* equivalent to the norm $\|\cdot\|_\mathcal{H}$ and where $\mathcal{H}$ endowed with $\|\cdot\|$ is hence *not* a Hilbert space.

Our aim is to prove convergence of $u_n := u_{h_n}^{\varepsilon_n}$ to a minimizer u^* in an appropriate sense. We first need to specify the assumptions on the choice of penalty and discretization parameters ε_n and h_n. For the penalty parameter, we assume that $(\varepsilon_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}_{>0}$ is an arbitrary zero sequence $\varepsilon_n \to 0$. Concerning the discretization, we assume that $(X_h)_{h \in I}$ with $I \subseteq \mathbb{R}_{>0}$ is a monotone family of finite dimensional subspaces of $\mathcal{H}$ with $\overline{\bigcup_{h>0} X_h} = \mathcal{H}$. An arbitrary zero sequence $(h_n)_{n \in \mathbb{N}} \subseteq I$ then satisfies the conditions (3). The corresponding discrete admissible sets are denoted by $\mathcal{A}_h := \mathcal{A} \cap X_h$.

With these notions and assumptions in hand, we show convergence $u_n \to u^*$:

Theorem 1 (Convergence as $(h, \varepsilon) \to (0,0)$)**.** *Let $(h_n)_{n \in \mathbb{N}}$ and $(\varepsilon_n)_{n \in \mathbb{N}}$ be arbitrary positive zero sequences. The sequence of minimizers $u_n := u_{h_n}^{\varepsilon_n}$ of $(M_{h_n}^{\varepsilon_n})$ satisfies convergence in the following sense: Any subsequence $(u_{n_k})_{k \in \mathbb{N}}$ contains a convergent subsequence $\left(u_{n_{k_\ell}} \right)_{\ell \in \mathbb{N}}$ whose limit is a minimizer u^* of the continuous constrained problem (M). Convergence holds with respect to both the weak topology of $\mathcal{H}$ and the topology induced by the semi-norm $\| \cdot \|$, i.e.*

$$u_{n_{k_\ell}} \rightharpoonup u^* \qquad and \qquad \| u_{n_{k_\ell}} - u^* \| \to 0. \tag{4}$$

Moreover, the entire sequence of energies converges:

$$e(u_n) \to e(u^*) \quad as \ well \ as \quad e_{\varepsilon_n}(u_n) \to e(u^*). \tag{5}$$

If the minimizer u^* is uniquely determined, we do not only have convergence of the energy sequences (5), but also convergence of the full sequence, i.e. (4) holds with $u_{n_{k_\ell}}$ replaced by u_n.

The remainder of the section is concerned with the proof of Theorem 1. It consists of four parts. As a first observation, we state the stability of the penalty method. By this we mean that given some sequence $(u_n)_{n \in \mathbb{N}} \subseteq \mathcal{H}$ and some zero sequence $\varepsilon_n \to 0$ such that the energies $e_{\varepsilon_n}(u_n)$ are uniformly bounded, a weak limit u must satisfy the admissibility condition $u \in \mathcal{A}$.

Proposition 2 (Stability of the penalty method)**.** *Let $\varepsilon_n \to 0$ be a non-negative zero sequence. Let $(u_n)_{n \in \mathbb{N}} \subseteq \mathcal{H}$ satisfy*

$$\sup_{n \in \mathbb{N}} e_{\varepsilon_n}(u_n) < \infty \tag{6}$$

and

$$u_n \rightharpoonup u_\infty$$

for some $u_\infty \in \mathcal{H}$. Then, $\lim_{n \to \infty} \|(g(u_n))_+\|_{L^2(\Omega)}^2 = 0$ and $u_\infty \in \mathcal{A}$, i.e. the weak limit u_∞ satisfies the constraint

$$g(u_\infty) \le 0.$$

Proof. By convexity of the functions $\| \cdot \|_{L^2(\Omega)}^2$, $(\cdot)_+$, and g, the function

$$u \mapsto \|(g(u))_+\|_{L^2(\Omega)}^2$$

is convex. Together with continuity we deduce that it is weakly lower semi-continuous.

¿From the weak convergence $u_n \rightharpoonup u_\infty$ we thus obtain

$$\|(g(u_\infty))_+\|_{L^2(\Omega)}^2 \le \liminf_{n \in \mathbb{N}} \|(g(u_n))_+\|_{L^2(\Omega)}^2. \tag{7}$$

¿From (6) we obtain the upper bound

$$\limsup_{n \in \mathbb{N}} e(u_n) \le \limsup_{n \in \mathbb{N}} e_{\varepsilon_n}(u_n) < \infty. \tag{8}$$

The weak lower semicontinuity of $e(\cdot)$—together with $u_n \rightharpoonup u_\infty$—yields the lower bound

$$e(u_\infty) \le \liminf_{n \in \mathbb{N}} e(u_n) \le \liminf_{n \in \mathbb{N}} e_{\varepsilon_n}(u_n) \le \sup_{n \in \mathbb{N}} e_{\varepsilon_n}(u_n) < \infty. \tag{9}$$

The combination of (8) and (9) shows that both sequences $e(u_n)$ and $e_{\varepsilon_n}(u_n)$ are bounded. Moreover, we have $e_{\varepsilon_n}(u_n) = e(u_n) + \frac{1}{2\varepsilon_n}\|(g(u_n))_+\|_{L^2(\Omega)}^2$ according to the definition of both energies. This produces

$$2\varepsilon_n(e_{\varepsilon_n}(u_n) - e(u_n)) = \|(g(u_n))_+\|_{L^2(\Omega)}^2.$$

The assumption $\varepsilon_n \to 0$ proves together with (7) that

$$\|(g(u_\infty))_+\|_{L^2(\Omega)}^2 \le \liminf_{n \in \mathbb{N}} \|(g(u_n))_+\|_{L^2(\Omega)}^2$$

$$\le \limsup_{n \in \mathbb{N}} \|(g(u_n))_+\|_{L^2(\Omega)}^2 = \limsup_{n \in \mathbb{N}} 2\varepsilon_n(e_{\varepsilon_n}(u_n) - e(u_n)) = 0,$$

i.e. $\lim_{n \to \infty} \|(g(u_n))_+\|_{L^2(\Omega)}^2 = 0$ and $u_\infty \in \mathcal{A}$. $\qquad\square$

Next, we analyze convergence for fixed $h \geq 0$ and $\varepsilon_n \to 0$.

Lemma 3 (Convergence as $\varepsilon \to 0$)**.** *Let $h \geq 0$ be fixed and let $(\varepsilon_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}_{>0}$ be an arbitrary positive zero sequence $\varepsilon_n \to 0$. The minimizers $u_h^{\varepsilon_n}$ of $(M_h^{\varepsilon_n})$ satisfy weak convergence with respect to the norm topology of $\mathcal{H}$ in the following sense: Any subsequence $\left(u_h^{\varepsilon_{n_k}} \right)_{k \in \mathbb{N}}$ contains a weakly convergent subsequence $\left(u_h^{\varepsilon_{n_{k_\ell}}} \right)_{\ell \in \mathbb{N}}$ whose limit is a minimizer u_h^0 of the constrained problem (M_h^0), i.e.*

$$u_h^{\varepsilon_{n_{k_\ell}}} \rightharpoonup u_h^0. \tag{10}$$

In addition, the energy sequence converges:

$$e(u_h^{\varepsilon_n}) \to e(u_h^0). \tag{11}$$

If the solution u_h^0 of the constrained problem is uniquely determined, we have weak convergence of the full sequence, i.e. (10) holds with $u_h^{\varepsilon_{n_{k_\ell}}}$ replaced by $u_h^{\varepsilon_n}$.

Proof. We may assume without loss of generality that $\varepsilon_n < 1$. Since $u_h^{\varepsilon_n}$ is the minimizer of the unconstrained problem $(M_h^{\varepsilon_n})$ and by definition of $e_{\varepsilon_n}(\cdot)$ we have

$$e_1(u_h^{\varepsilon_n}) \leq e_{\varepsilon_n}(u_h^{\varepsilon_n}) \leq e_{\varepsilon_n}(u_h^0) = e(u_h^0). \tag{12}$$

Therefore, the sequence of minimizers $u_h^{\varepsilon_n}$ has bounded energy. From assumption (A2) and in particular coercivity of $e_1(\cdot)$, we obtain boundedness of $u_h^{\varepsilon_n}$ in norm. This means that any subsequence $\left(u_h^{\varepsilon_{n_k}} \right)_{k \in \mathbb{N}}$ is bounded and must have a weakly convergent subsequence $u_h^{\varepsilon_{n_{k_\ell}}} \rightharpoonup u_h^*$. It remains to prove that first $u_h^* \in \mathcal{A}_h$ and that second it is indeed a minimizer of $e(\cdot)$.

The first statement, i.e. $u_h^* \in \mathcal{A}_h$, follows immediately from Proposition 2. From

$$e(u_h^{\varepsilon_{n_{k_\ell}}}) \leq e_{\varepsilon_{n_{k_\ell}}}(u_h^{\varepsilon_{n_{k_\ell}}}) \leq e(u_h^0)$$

and weak lower semicontinuity of $e(\cdot)$ we conclude

$$e(u_h^*) \leq \liminf_{\ell \in \mathbb{N}} e(u_h^{\varepsilon_{n_{k_\ell}}}) \leq \limsup_{\ell \in \mathbb{N}} e(u_h^{\varepsilon_{n_{k_\ell}}}) \leq e(u_h^0).$$

Since $u_h^0 \in \mathcal{A}_h$ is a minimizer of $e(\cdot)$, there must hold $e(u_h^*) = e(u_h^0)$, i.e. $u_h^* \in \mathcal{A}_h$ is a minimizer of (M_h^0) as well. Moreover, we obtain convergence in energy $e(u_h^{\varepsilon_{n_{k_\ell}}}) \to e(u_h^0)$.

In particular, this proves that each subsequence $e(u_h^{\varepsilon_{n_k}})$ contains a subsequence $e(u_h^{\varepsilon_{n_{k_\ell}}})$ which tends to the independent limit $e(u_h^0)$. Elementary calculus thus predicts that the entire energy sequence $e(u_h^{\varepsilon_n})$ converges to $e(u_h^0)$. Finally, if u_h^* is uniquely determined, we have seen that each subsequence of $(u_h^{\varepsilon_n})_{n \in \mathbb{N}}$ contains a subsequence whose weak limit is the unique u_h^*. From this one can already conclude $\lim_{n \to \infty} u_h^{\varepsilon_n} = u_h^*$ in the weak topology. $\square$

In a third step we use similar arguments to prove convergence for fixed $\varepsilon \geq 0$ as $h \to 0$.

Lemma 4 (Convergence as $h \to 0$)**.** *Let $\varepsilon \geq 0$ be fixed and let $(h_n)_{n \in \mathbb{N}} \subseteq I$ be an arbitrary zero sequence $h_n \to 0$. Then, the minimizers $u_{h_n}^{\varepsilon}$ of $(M_{h_n}^{\varepsilon})$ satisfy weak convergence with respect to the norm topology of $\mathcal{H}$ in the following sense: Any subsequence $\left(u_{h_{n_k}}^{\varepsilon} \right)_{k \in \mathbb{N}}$ contains a weakly convergent subsequence $\left(u_{h_{n_{k_\ell}}}^{\varepsilon} \right)_{\ell \in \mathbb{N}}$ whose limit is a minimizer u_0^{ε} of the continuous penalized problem (M_0^{ε}), i.e.*

$$u_{h_{n_{k_\ell}}}^{\varepsilon} \rightharpoonup u_0^{\varepsilon} \tag{13}$$

Moreover, the entire energy sequence converges:

$$e_{\varepsilon}(u_{h_n}^{\varepsilon}) \to e_{\varepsilon}(u_0^{\varepsilon}). \tag{14}$$

If the solution u_0^{ε} of the continuous problem is uniquely determined, we have weak convergence of the full sequence, i.e. (13) holds with $u_{h_{n_{k_\ell}}}^{\varepsilon}$ replaced by $u_{h_n}^{\varepsilon}$.

Proof. We assume without loss of generality $h_n \leq 1$. Then, from monotonicity of spaces we have $X_1 \subseteq X_{h_n}$ and hence

$$e_\varepsilon(u_{h_n}^\varepsilon) \leq e_\varepsilon(u_1^\varepsilon). \tag{15}$$

Put differently, the sequence $(u_{h_n}^\varepsilon)_{n\in\mathbb{N}}$ has bounded energy. From coercivity (A2) we deduce boundedness of the sequence in $\mathcal{H}$. In particular, each subsequence $\left(u_{h_{n_k}}^\varepsilon\right)_{k\in\mathbb{N}}$ is also bounded and has therefore a weakly convergent subsequence

$$u_{h_{n_{k_\ell}}}^\varepsilon \rightharpoonup u_*^\varepsilon.$$

It remains to prove that u_*^ε is a minimizer.

Let u_0^ε denote a minimizer of the continuous problem (M_0^ε). ¿From $\overline{\bigcup_{h>0} X_h} = \mathcal{H}$, we know that for each $\delta > 0$ there is an integer $L \in \mathbb{N}$ such that for all $\ell \geq L$ there exists some $\widetilde{u}_{h_{n_{k_\ell}}}^\varepsilon \in X_{h_{n_{k_\ell}}}$ with

$$\|u_0^\varepsilon - \widetilde{u}_{h_{n_{k_\ell}}}^\varepsilon\|_{\mathcal{H}} \leq \delta.$$

Together with continuity of $e_\varepsilon(\cdot)$, we know that for arbitrary $\eta > 0$ there exists some index $L \in \mathbb{N}$ such that for all $\ell \geq L$ we may thus choose $\widetilde{u}_{h_{n_{k_\ell}}}^\varepsilon \in X_{h_{n_{k_\ell}}}$ with

$$e_\varepsilon(\widetilde{u}_{h_{n_{k_\ell}}}^\varepsilon) \leq e_\varepsilon(u_0^\varepsilon) + \eta.$$

Recall that $u_{h_{n_{k_\ell}}}^\varepsilon$ is a minimizer and therefore $e_\varepsilon(u_{h_{n_{k_\ell}}}^\varepsilon) \leq e_\varepsilon(\widetilde{u}_{h_{n_{k_\ell}}}^\varepsilon)$. Since this holds for all $\eta > 0$, and together with (weak lower semi-)continuity of $e_\varepsilon(\cdot)$, we conclude

$$e_\varepsilon(u_*^\varepsilon) \leq \liminf_{\ell\to\infty} e_\varepsilon(u_{h_{n_{k_\ell}}}^\varepsilon) \leq e_\varepsilon(u_0^\varepsilon).$$

This means that u_*^ε is in fact a minimizer. Moreover, the preceding estimates yield $\liminf_{\ell\in\mathbb{N}} e_\varepsilon(u_{h_{n_{k_\ell}}}^\varepsilon) = e_\varepsilon(u_0^\varepsilon)$. By extracting an additional subsequence, we may thus generate $\lim_{\ell\in\mathbb{N}} e_\varepsilon(u_{h_{n_{k_\ell}}}^\varepsilon) = e_\varepsilon(u_0^\varepsilon)$. Arguing as in Lemma 3, we may conclude convergence of the energy sequence as well as weak convergence of the full sequence provided that the limit is unique. $\square$

So far, we have proven (all limits are with respect to the weak topology)

$$\lim_{h\to 0} \lim_{\varepsilon\to 0} u_h^\varepsilon = u^* = \lim_{\varepsilon\to 0} \lim_{h\to 0} u_h^\varepsilon;$$

we next prove the statement of Theorem 1. We stress that no assumptions on the regularity of the analytical solution are necessary; no knowledge about existence or estimates for Lagrange multipliers are used.

Proof of Theorem 1. Let $u_0^{\varepsilon_n}$ denote a minimizer of the continuous and penalized problem $(M_0^{\varepsilon_n})$ and let $u_{h_n}^0$ denote a minimizer of the discrete constrained problem $(M_{h_n}^0)$. We observe

$$e(u_0^0) \leq e(u_{h_n}^{\varepsilon_n}) \leq e_{\varepsilon_n}(u_{h_n}^{\varepsilon_n}) \leq e_{\varepsilon_n}(0) = e(0) = 0. \tag{16}$$

Therefore the sequence $u_n := u_{h_n}^{\varepsilon_n}$ is bounded and every subsequence $(u_{n_k})_{k\in\mathbb{N}}$ has a weakly convergent subsequence $u_{n_{k_\ell}} \rightharpoonup \widetilde{u}^*$. From Lemma 4 we obtain the existence of a subsequence, which we write for simplicity as $(u_{h_s}^{\varepsilon_s})_{s\in\mathbb{N}}$, such that the associated sequence $(u_{h_s}^0)_{s\in\mathbb{N}}$ converges weakly to a minimizer u^* and $e(u_{h_s}^0) \to e(u^*)$. Hence,

$$e(u_{h_s}^{\varepsilon_s}) \leq e_{\varepsilon_s}(u_{h_s}^{\varepsilon_s}) \leq e_{\varepsilon_s}(u_{h_s}^0) = e(u_{h_s}^0) \to e(u^*).$$

¿From this it follows, again by weak lower semicontinuity of $e(\cdot)$, that $e(\widetilde{u}^*) \leq e(u^*)$. Recalling $e_{\varepsilon_s}(u_{h_s}^{\varepsilon_s}) \leq e_{\varepsilon_s}(u_{h_0}^0) = e(u_{h_0}^0)$, we see that $u_{h_s}^{\varepsilon_s}$ satisfies the assumptions of Proposition 2, and therefore $\widetilde{u}^* \in \mathcal{A}$.

6

Finally we obtain that $\widetilde{u}^*$ is a minimizer, i.e. the desired convergence result with respect to the weak topology. Moreover,

$$e(u^*) = e(\widetilde{u}^*) \le \liminf_{s \in \mathbb{N}} e(u_{h_s}^{\varepsilon_s}) \le \limsup_{s \in \mathbb{N}} e(u_{h_s}^{\varepsilon_s}) \le \limsup_{s \in \mathbb{N}} e(u_{h_s}^0) = e(u^*)$$

proves convergence of the energy $e(u_{h_s}^{\varepsilon_s}) \to e(u^*)$. The same argument with $e(\cdot)$ replaced by $e_{\varepsilon_s}(\cdot)$ yields $e_{\varepsilon_s}(u_{h_s}^{\varepsilon_s}) \to e(u^*)$. Arguing as above, we may even derive energy convergence $e(u_n) \to e(u^*)$ as well as $e_{\varepsilon_n}(u_n) \to e(u^*)$.

The strong convergence in the energy semi-norm $\|\!|\cdot|\!\|$ is obtained by a bootstrapping argument: Recall the definition of the energy

$$e(u) = \frac{1}{2}\|\!|u|\!\|^2 - \Phi(u).$$

In other words, the energy semi-norm satisfies

$$\|\!|u|\!\|^2 = 2e(u) + 2\Phi(u).$$

¿From the construction above, we have $u_s := u_{h_s}^{\varepsilon_s} \rightharpoonup \widetilde{u}^*$ and $e(u_s) \to e(\widetilde{u}^*)$. Recall that $\Phi(u)$ is linear and continuous with respect to the $\mathcal{H}$-norm so that that $\Phi(u_s) \to \Phi(\widetilde{u}^*)$. Together we see

$$\|\!|u_s|\!\|^2 = 2e(u_s) + 2\Phi(u_s) \longrightarrow 2e(\widetilde{u}^*) + 2\Phi(\widetilde{u}^*) = \|\!|\widetilde{u}^*|\!\|^2.$$

Since $\|\!|\cdot|\!\|$ stems from a continuous semi-scalar product we immediately conclude

$$\begin{aligned}\|\!|\widetilde{u}^* - u_s|\!\|^2 = \langle\!\langle \widetilde{u}^* - u_s \,,\, \widetilde{u}^* - u_s \rangle\!\rangle &= \|\!|\widetilde{u}^*|\!\|^2 - 2\langle\!\langle \widetilde{u}^* \,,\, u_s \rangle\!\rangle + \|\!|u_s|\!\|^2 \\ &\longrightarrow \|\!|\widetilde{u}^*|\!\|^2 - 2\|\!|\widetilde{u}^*|\!\|^2 + \|\!|\widetilde{u}^*|\!\|^2 = 0,\end{aligned}$$

which means convergence in the energy semi-norm. $\qquad\square$

3 Thin-film micromagnetics

For many applications in stationary micromagnetics, the model due to Landau and Lifshits [18] is nowadays accepted as relevant. From a computational point of view it is highly challenging. Different length scales involved make large samples of several μm in diameter hardly accessible for direct calculations. Therefore, various reduced models have been proposed and analyzed that cover certain asymptotic regimes, see [9] for an overview.

Here, we discuss the reduced model proposed in [7]. It covers the regime of very thin but relatively large ferromagnetic samples. In [8] the authors obtain the reduced model for vanishing thickness of the sample—under certain assumptions—as Γ-limit of the full problem due to Landau and Lifshits.

Micromagnetic thin-film problem ($\boldsymbol{TF}$): Let $\Omega \subseteq \mathbb{R}^2$ be a bounded Lipschitz domain that represents the ferromagnetic sample, whose thickness is neglected by the model. With an in-plane applied exterior field $\mathbf{f} : \Omega \to \mathbb{R}^2$, we seek a magnetization $\mathbf{m}^* : \Omega \to \mathbb{R}^2$ that satisfies the convex pointwise constraint $g(\mathbf{m}) = |\mathbf{m}| - 1 \le 0$ almost everywhere and minimizes the reduced energy

$$e(\mathbf{m}) = \frac{1}{2}\int_{\mathbb{R}^3} |\nabla p|^2 \, dx + \frac{q}{2}\int_\Omega \mathbf{m}_2^2 \, dx - \int_\Omega \mathbf{f} \cdot \mathbf{m} \, dx. \tag{17}$$

The magnetic potential $p : \mathbb{R}^3 \to \mathbb{R}$ is determined by the reduced magnetostatic Maxwell equation

$$\int_{\mathbb{R}^3} \nabla p \cdot \nabla v \, dx = \int_\Omega \mathbf{m} \cdot \nabla v(x, 0) \, dx \quad \text{for all} \quad v \in \mathcal{D}(\mathbb{R}^3) := \mathcal{C}_c^\infty(\mathbb{R}^3), \tag{18}$$

stated here in a distributional sense. The material dependent parameter $q > 0$ measures the strength of the uniaxial crystalline anisotropy. For soft materials such as permalloy, where $q \ll 1$, one usually drops this energy contribution by setting $q = 0$.

3.1. The magnetostatic Maxwell equation. The magnetic potential is the solution of the variational formulation (18). Understanding the potential p is crucial to define the appropriate function space for the magnetization $\mathbf{m}$. We denote by $\mathbf{n}$ the outer normal in $\mathbb{R}^2$ of Ω and by $[\cdot]$ the jump across Ω. For smooth $\mathbf{m} \in \mathcal{C}^1(\overline{\Omega})$, every weak solution p of (18) which is sufficiently smooth, i.e. $p \in \mathcal{C}^2(\mathbb{R}^3 \setminus \overline{\Omega}) \cap \mathcal{C}(\mathbb{R}^3) \cap \mathcal{C}^1(\mathbb{R}^2 \times \mathbb{R}_{\geq 0}) \cap \mathcal{C}^1(\mathbb{R}^2 \times \mathbb{R}_{\leq 0})$, solves the strong form

$$
\begin{aligned}
\Delta p &= 0 &&\text{in} \quad \mathbb{R}^3 \setminus \overline{\Omega}, \\
\left[\frac{\partial p}{\partial x_3}\right] &= \nabla \cdot \mathbf{m} &&\text{on} \quad \Omega, \\
[p] &= 0 &&\text{on} \quad \Omega, \\
\mathbf{m} \cdot \mathbf{n} &= 0 &&\text{on} \quad \Gamma = \partial\Omega \subseteq \mathbb{R}^2.
\end{aligned}
\tag{19}
$$

We give a brief definition of some Sobolev spaces needed in the following. The reader is referred to e.g. [1] for a detailed discussion. The Sobolev space $H^1(G)$ for some Lipschitz domain G is defined in the usual way

$$
H^1(G) = \{u \in L^2(G) \,|\, \nabla u \in L^2(G) \text{ in a weak sense}\}
$$

and equipped with the natural norm $\|u\|_{H^1(G)}^2 = \|u\|_{L^2(G)}^2 + \|\nabla u\|_{L^2(G)}^2$. We then define the fractional order Sobolev space

$$
H^{1/2}(\Omega) = \{u \in L^2(\Omega) \,|\, \text{there is some } \widetilde{u} \in H^1(\mathbb{R}^3) \text{ with } \widetilde{u}|_\Omega = u\}
$$

as the space of traces of H^1 functions restricted to Ω. This space may be equipped with the norm

$$
\|u\|_{H^{1/2}(\Omega)} = \inf_{\substack{\widetilde{u} \in H^1(\mathbb{R}^3) \\ \widetilde{u}|_\Omega = u}} \|\widetilde{u}\|_{H^1(\mathbb{R}^3)},
$$

which yields a Hilbert space. Finally its dual space with respect to the extended L^2 scalar product is denoted by $\widetilde{H}^{-1/2}(\Omega)$.

Recall the definition of the simple-layer potential

$$
S\varphi(x) = \frac{1}{4\pi} \int_\Omega \frac{\varphi(y)}{|x-y|} \, dy \quad x \in \mathbb{R}^3 \setminus \overline{\Omega}
\tag{20}
$$

associated with the Laplace-operator in 3D. The operator S can be extended continuously to $S \in L(\widetilde{H}^{-1/2}(\Omega); H^1_{\ell oc}(\mathbb{R}^3))$ and its trace

$$
V\varphi(x) = \frac{1}{4\pi} \int_\Omega \frac{\varphi(y)}{|x-y|} \, dy \quad \text{for } x \in \Omega
$$

can be extended continuously to $V \in L(\widetilde{H}^{-1/2}(\Omega); H^{1/2}(\Omega))$, see [25]. Moreover, the operator V satisfies the ellipticity estimate

$$
\|\varphi\|_{\widetilde{H}^{-1/2}(\Omega)}^2 \leq C_{ell} \langle \varphi, \, V\varphi \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)}
$$

with some constant $C_{ell} > 0$ that depends only on Ω. Altogether, we have that

$$
(\varphi, \psi)_V := \langle \varphi, \, V\varphi \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)}
\tag{21}
$$

defines an equivalent scalar product on $\widetilde{H}^{-1/2}(\Omega)$.

Proposition 5 (Jump conditions of the simple-layer potential). *Given some $\varphi \in \widetilde{H}^{-1/2}(\Omega)$, the simple-layer potential $S\varphi$ satisfies*

$$
[S\varphi] = 0 \in H^{1/2}(\Omega) \quad \text{and} \quad \left[\frac{\partial S\varphi}{\partial x_3}\right] = -\varphi \in \widetilde{H}^{-1/2}(\Omega). \quad \square
$$

Remark. *The proof of Proposition 5 can be found e.g. in [24, Thm. 3.3.1] for the simple-layer potential defined on a closed surface in $\mathbb{R}^3$. Whereas the proof of the continuity follows literally, the proof of the jump relation for the normal derivative needs some modifications for the present case of a screen. We use similar ideas as in [25]: Choose some bounded Lipschitz domain G_- such that $\Omega \subseteq \partial G_-$ and such that the normal vector $\mathbf{n}$ of G_- satisfies $\mathbf{n}|_\Omega = (0,0,1)$. Given some $\varphi \in \widetilde{H}^{-1/2}(\Omega)$, the second Green's formula for $u := S\varphi$ in the interior and exterior domains G_- and $G_+ := \mathbb{R}^3 \backslash \overline{G}_-$ and the same arguments as in the proof of [24, Thm. 3.3.1] yield the statement.* $\qquad\square$

Corollary 6. *With the simple-layer potential S from (20) and given $\nabla \cdot \mathbf{m} \in \widetilde{H}^{-1/2}(\Omega)$, the function $S(-\nabla \cdot \mathbf{m})$ is a solution to the Maxwell equation (18).* $\qquad\square$

Remark. *First, we have shown that the regularity assumption $\nabla \cdot \mathbf{m} \in \widetilde{H}^{-1/2}(\Omega)$, ensures existence of p. Second, we want to comment on the constraint $\mathbf{m} \cdot \mathbf{n} = 0$ on Γ that arises for smooth solutions. Suppose $\mathbf{m} \cdot \mathbf{n} \neq 0$ on Γ, then a solution to the reduced Maxwell equation (18) does not in general satisfy $p \in H^1_{loc}(\mathbb{R}^3)$. Assume e.g. $\mathbf{m} = (1,0)^T$ constant. Then, $\mathbf{m} \cdot \mathbf{n} \neq 0 \in L^2(\Gamma)$ and $\nabla \cdot \mathbf{m} = 0 \in L^2(\Omega)$. Choose some smooth and bounded domain $G \subseteq \mathbb{R}^3$ with $\overline{\Omega} \subseteq G$. Integration by parts in the Maxwell equation (18) yields for all $v \in \mathcal{D}(G)$*

$$\int_G \nabla p \cdot \nabla v \, dX = \int_\Omega \mathbf{m} \cdot \nabla v \, dx = -\int_\Omega \nabla \cdot \mathbf{m} v \, dx + \int_\Gamma (\mathbf{m} \cdot \mathbf{n}) v|_\Gamma \, ds. \tag{22}$$

On the right-hand side of (22), the integral over Ω vanishes due to the choice of $\mathbf{m}$. The functional $v \mapsto \int_\Gamma (\mathbf{m} \cdot \mathbf{n}) v \, ds$ cannot be extended to a linear functional on $H^1(G)$ since the restriction to Γ is not well-defined. We conclude that $v \mapsto (\nabla p, \nabla v)_\Omega$ does not define a continuous functional in $H^1(G)$. This means $\nabla p \notin L^2(G)$ and, therefore, $p \notin H^1(G)$. $\qquad\square$

3.2. Energy space. A function $v \in L^2(\Omega)$ is called a weak divergence of $\mathbf{m} \in L^2(\Omega)^2$ if it satisfies

$$\int_\Omega v\varphi \, dx = -\int_\Omega \mathbf{m} \cdot \nabla\varphi \, dx \quad \text{for all} \quad \varphi \in \mathcal{D}(\Omega). \tag{23}$$

In this case, and according to the fundamental theorem of calculus of variations, the weak divergence of $\mathbf{m}$ is unique, and we simply write $\nabla \cdot \mathbf{m} := v$.

We define the space

$$H^1(\nabla\cdot; \Omega) := \{\mathbf{m} \in L^2(\Omega)^2 \,|\, \nabla \cdot \mathbf{m} \in L^2(\Omega)\} \tag{24}$$

with the canonical norm $\|\mathbf{m}\|_{H^1(\nabla\cdot;\Omega)} := (\|\mathbf{m}\|^2_{L^2(\Omega)^2} + \|\nabla \cdot \mathbf{m}\|^2_{L^2(\Omega)})^{1/2}$. We further define $H^1_0(\nabla\cdot;\Omega) := \overline{\mathcal{D}(\Omega)^2}^{\|\cdot\|_{H^1(\nabla\cdot;\Omega)}}$.

Proposition 7 ([13, Theorem 2.5 and Theorem 2.6]). *The mapping $f_n : \mathbf{v} \mapsto \mathbf{v} \cdot \mathbf{n}|_\Gamma$ defined on $\mathcal{D}(\overline{\Omega})^2$ can be extended by continuity to a linear and continuous mapping from $H^1(\nabla\cdot;\Omega)$ into $H^{-1/2}(\Gamma)$. There holds $H^1_0(\nabla\cdot;\Omega) = \ker(f_n) = \{\mathbf{m} \in H^1(\nabla\cdot;\Omega) \,|\, \mathbf{m} \cdot \mathbf{n} = 0\}$, and $\mathcal{D}(\Omega)^2$ is dense in $H^1_0(\nabla\cdot;\Omega)$.* $\qquad\square$

We have to do one last step to define the appropriate space for the magnetization. Namely, we have to permit $\nabla \cdot \mathbf{m} \in \widetilde{H}^{-1/2}(\Omega) \supseteq L^2(\Omega)$.

We define the energy space for the magnetization

$$\mathcal{H} = \overline{H^1_0(\nabla\cdot;\Omega)}^{\|\cdot\|_\mathcal{H}} \tag{25}$$

where

$$\|\mathbf{m}\|^2_\mathcal{H} := \|\mathbf{m}\|^2_{L^2(\Omega)^2} + \|\nabla \cdot \mathbf{m}\|^2_{\widetilde{H}^{-1/2}(\Omega)}. \tag{26}$$

By construction $\mathcal{H}$ is a Hilbert space and $\mathcal{D}(\Omega)^2$ is a dense subspace.

Finally, we make a last remark on $\mathcal{H}$ before analyzing the existence of minimizers $\mathbf{m}^*$ in our function setting.

Lemma 8. *For all functions* $\mathbf{m} \in \mathcal{H}$, *we have* $\langle \nabla \cdot \mathbf{m} \, , \, 1 \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)} = 0$.

Proof. According to the Gauss divergence theorem, there holds for all $\mathbf{m} \in \mathcal{D}(\Omega)^2$

$$\int_\Omega \nabla \cdot \mathbf{m} \, dx = \int_\Gamma \mathbf{m} \cdot \mathbf{n} \, ds = 0,$$

where the right-hand side vanishes in our case. This implies for all $\mathbf{m} \in \mathcal{D}(\Omega)^2$

$$(\nabla \cdot \mathbf{m}, 1)_{L^2(\Omega)} = \langle \nabla \cdot \mathbf{m} \, , \, 1 \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)} = 0.$$

Due to the density of $\mathcal{D}(\Omega)^2$ in $\mathcal{H}$, the statement immediately follows by continuity. $\qquad\square$

Given $\mathbf{m} \in \mathcal{H}$, we may represent p as the simple-layer potential of $\nabla \cdot \mathbf{m} \in \widetilde{H}^{-1/2}(\Omega)$. From the mapping properties of the simple-layer potential, we immediately conclude $p \in H^1_{\ell oc}(\mathbb{R}^3) \supseteq \text{range}(S)$. However, $H^1_{\ell oc}(\mathbb{R}^3)$ is not a normed space. Also, the fact $\nabla p \in L^2(\mathbb{R}^3)^3$ implies further regularity and is necessary for the energy $e(\mathbf{m})$ defined in (17) to be finite. We now establish an appropriate Hilbert space for the magnetostatic potential p.

We define the set

$$\widetilde{B}_1^2(\mathbb{R}^3) := \{ u \in H^1_{\ell oc}(\mathbb{R}^3) \, | \, \nabla u \in L^2(\mathbb{R}^3)^3 \} \tag{27}$$

associated with the semi-norm $\|u\|_{B_1^2(\mathbb{R}^3)} := \|\nabla u\|_{L^2(\mathbb{R}^3)^3}$. Further, we define the Beppo-Levi space

$$B_1^2(\mathbb{R}^3) := \widetilde{B}_1^2(\mathbb{R}^3)/\mathbb{R} \tag{28}$$

by factoring out the constant functions. Note that $\| \cdot \|_{B_1^2(\mathbb{R}^3)}$ now in fact is a norm.

Proposition 9 ([6, Corollaire 1.1, Théorème 2.1]). $B_1^2(\mathbb{R}^3)$ *is a Hilbert space, and* $\mathcal{D}(\mathbb{R}^3)$ *is a dense subspace of* $B_1^2(\mathbb{R}^3)$. $\qquad\square$

Lemma 10. *There is a continuous linear lifting operator* $L : H^{1/2}(\Omega) \to B_1^2(\mathbb{R}^3)$, *i.e. for* $v \in H^{1/2}(\Omega)$ *we have*

$$v = (Lv)|_\omega \quad \text{and} \quad \|\nabla(Lv)\|_{L^2(\mathbb{R}^3)^3} \leq C\|v\|_{H^{1/2}(\Omega)}$$

with $C > 0$ *the operator norm of* L.

Proof. First, choose some bounded Lipschitz domain $G \subseteq (\mathbb{R}^3)$ with $\Omega \subseteq \partial G$. Let $v \in H^{1/2}(\Omega)$ be given. Then, e.g. from [20, Theorem 3.37], we conclude existence of some extension $\widetilde{v} \in H^1(G)$ with $\|\widetilde{v}\|_{H^1(G)} \lesssim \|v\|_{H^{1/2}(\Omega)}$. The results of [6, Théorème 8.1] provide the existence of a linear and continuous lifting operator from $H^1(G)$ to $B_1^2(\mathbb{R}^3)$, which finally proves the full linear and continuous extension. $\quad\square$

3.3. Well-posedness of the micromagnetic thin-film problem. We proceed to analyze well-posedness of the problem (TF) in our functional analytic setting. As a first observation, in Proposition 12, we prove that for any given $\mathbf{m} \in \mathcal{H}$ there exists a uniquely determined magnetostatic potential $p \in B_1^2(\mathbb{R}^3)$. Moreover, we will see that $\int_{\mathbb{R}^3} |\nabla p|^2 \, dx = \|\nabla \cdot \mathbf{m}\|_V^2$. This reveals that we can restate our energy functional from (17)–(18) in the abstract form

$$e(\mathbf{m}) = \frac{1}{2}\|\mathbf{m}\|^2 - \Phi(\mathbf{m}) \tag{29}$$

of Section 1 with energy semi-norm and linear functional

$$\|\mathbf{m}\|^2 := \|\nabla \cdot \mathbf{m}\|_V^2 + q\|\mathbf{m}_2\|_{L^2(\Omega)}^2 \quad \text{and} \quad \Phi(\mathbf{m}) = (\mathbf{f}, \mathbf{m})_{L^2(\Omega)}. \tag{30}$$

The outline of the remaining section now reads as follows: In Lemma 13, we prove that the energy $e(\mathbf{m})$ and the penalized energy

$$e_\varepsilon(\mathbf{m}) = e(\mathbf{m}) + \frac{1}{2\varepsilon}\|(|\mathbf{m}| - 1)_+\|_{L^2(\Omega)}^2 \tag{31}$$

satisfy the coercivity assumptions (A1)–(A2). Lemma 14 states that the energy semi-norm is in fact positive definite if $q > 0$. This yields not only existence, but even uniqueness of the minimizer $\mathbf{m}^*$. Finally, Lemma 16 uses the corresponding variational inequality to prove continuous dependence of $\mathbf{m}^*$, on the data.

We collect the main results of this section in the following theorem:

Theorem 11. *The following statements on the micromagnetic thin-film problem* (TF) *are true:*

(i) *The micromagnetic thin-film problem fits into the abstract setting of Section 1 with the definitions of* (30). *The energy functionals* $e(\mathbf{m})$ *and* $e_\varepsilon(\mathbf{m}) = e(\mathbf{m}) + \frac{1}{2\varepsilon}\|(|\mathbf{m}|-1)_+\|^2_{L^2(\Omega)}$ *satisfy the coercivity assumptions* (A1)–(A2).

(ii) *The micromagnetic thin-film problem has a solution* $\mathbf{m}^* \in \mathcal{A}$. *For* $q > 0$, $\|\cdot\|$ *is a norm and hence the minimizer* $\mathbf{m}^*$ *is unique. However,* $\mathcal{H}$ *endowed with the norm* $\|\cdot\|$ *is* not *complete.*

(iii) *If* $\mathbf{m}^1$ *and* $\mathbf{m}^2$ *are solutions of* (TF) *for applied fields* $\mathbf{f}^1$ *and* $\mathbf{f}^2$, *respectively, then*

$$\|\mathbf{m}^1 - \mathbf{m}^2\| \leq \sqrt{2}\|\mathbf{f}^1 - \mathbf{f}^2\|^{1/2}_{L^1(\Omega)}, \tag{32}$$

i.e. $\mathbf{m}^*$ *depends Hölder continuously on the applied field* $\mathbf{f}$. $\qquad\qquad\square$

We recall the variational formulation

$$\int_{\mathbb{R}^3} \nabla p \cdot \nabla v \, dx = -\int_\omega \nabla \cdot \mathbf{m} \, v \, dx \quad \text{for all} \quad v \in \mathcal{D}(\mathbb{R}^3). \tag{33}$$

of the magnetostatic Maxwell equation. As discussed above, $\mathbf{m} \in \mathcal{H}$ satisfies all constraints and the necessary regularity. Therefore, (33) may be stated as

$$(\nabla p, \nabla v)_{L^2(\mathbb{R}^3)^3} = -\langle \nabla \cdot \mathbf{m}, v \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)} \quad \text{for all} \quad v \in \mathcal{D}(\mathbb{R}^3) \tag{34}$$

in our functional setting.

Proposition 12. *The following statements are true:*

(i) *Given* $\mathbf{m} \in \mathcal{H}$, *there is a uniquely determined* $p \in B_1^2(\mathbb{R}^3)$ *with* (34).

(ii) *Equation* (34) *holds with* $\mathcal{D}(\mathbb{R}^3)$ *replaced with the full space* $B_1^2(\mathbb{R}^3)$.

(iii) *The mapping* $\mathcal{P} : \mathcal{H} \to L^2(\mathbb{R}^3)^3$ *that maps* $\mathbf{m}$ *to the corresponding stray field* $\mathcal{P}(\mathbf{m}) := -\nabla p$ *is a linear and continuous operator.*

(iv) *For* $\mathbf{m}, \widetilde{\mathbf{m}} \in \mathcal{H}$, *there holds* $\langle \mathcal{P}\mathbf{m}, \mathcal{P}\widetilde{\mathbf{m}} \rangle_{L^2(\mathbb{R}^3)^3} = \langle \nabla \cdot \mathbf{m}, V(\nabla \cdot \widetilde{\mathbf{m}}) \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)}$.

(v) *In particular, there holds* $\|\mathcal{P}\mathbf{m}\|^2_{L^2(\mathbb{R}^3)^3} = \|\nabla \cdot \mathbf{m}\|^2_V \simeq \|\nabla \cdot \mathbf{m}\|^2_{\widetilde{H}^{-1/2}(\Omega)}$.

Proof. Let $\mathbf{m} \in \mathcal{H}$ be fixed. We first consider $F_\mathbf{m}(v) := \langle \nabla \cdot \mathbf{m}, v \rangle_{\widetilde{H}^{-1/2}(\omega) \times H^{1/2}(\omega)}$ for arbitrary $v \in \mathcal{D}(\mathbb{R}^3)$. According to Lemma 8, we have

$$\langle \nabla \cdot \mathbf{m}, v \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)} = \langle \nabla \cdot \mathbf{m}, v - \lambda \rangle \quad \text{for all constants} \quad \lambda \in \mathbb{R}.$$

We consider the cylindrical domain $\widehat{\Omega} := \Omega \times [0, 1]$. With $v \in \mathcal{D}(\mathbb{R}^3)$ and $\lambda := (1/|\widehat{\Omega}|) \int_{\widehat{\Omega}} v \, dx$, the continuity of the trace operator and a Poincaré inequality show

$$\|v - \lambda\|_{H^{1/2}(\Omega)} \lesssim \|v - \lambda\|_{H^1(\widehat{\Omega})} \lesssim \|\nabla v\|_{L^2(\widehat{\Omega})^3} \leq \|\nabla v\|_{L^2(\mathbb{R}^3)^3}.$$

This proves that $F_\mathbf{m}$ defines a linear and continuous functional $F_\mathbf{m} : \mathcal{D}(\mathbb{R}^3) \to \mathbb{R}$ with respect to $\|\cdot\|_{B_1^2(\mathbb{R}^3)}$ and operator norm $\|F_\mathbf{m}\| \leq \|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\Omega)}$. Since $\mathcal{D}(\mathbb{R}^3)$ is dense in $B_1^2(\mathbb{R}^3)$, the functional $F_\mathbf{m}$ may be extended continuously to the entire Beppo-Levi space while preserving the operator norm. Since the left-hand side of (34) is the scalar product of $B_1^2(\mathbb{R}^3)$, the variational formulation may be extended to the full space $B_1^2(\mathbb{R}^3)$. This proves (ii).

The Riesz representation theorem provides existence and uniqueness of a solution $p \in B_1^2(\mathbb{R}^3)$, since (34) may be written as

$$(p, v)_{B_1^2(\mathbb{R}^3)} = F_{\mathbf{m}}(v) \quad \text{for all} \quad v \in B_1^2(\mathbb{R}^3),$$

which yields statement (i).

The Riesz theorem furthermore implies

$$\|\mathcal{P}\mathbf{m}\|_{L^2(\mathbb{R}^3)^3} = \|p\|_{B_1^2(\mathbb{R}^3)} = \|F_{\mathbf{m}}\| \lesssim \|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\omega)}.$$

In particular the mapping $\mathcal{P} : \mathcal{H} \to L^2(\mathbb{R}^3)^3$ is well defined and continuous. Linearity follows from the composition $\mathcal{P} : \mathbf{m} \mapsto \nabla \cdot \mathbf{m} \mapsto p \mapsto -\nabla p$, which finally proves (iii).

Now, we prove the converse estimate $\|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\omega)} \lesssim \|F_{\mathbf{m}}\|$. Lemma 10 and (34) imply

$$\frac{|\langle \nabla \cdot \mathbf{m}, v \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)}|}{\|v\|_{H^{1/2}(\Omega)}} \lesssim \frac{|\langle \mathcal{P}\mathbf{m}, \nabla(Lv) \rangle_{L^2(\mathbb{R}^3)^3}|}{\|\nabla(Lv)\|_{L^2(\mathbb{R}^3)^3}} \leq \|\mathcal{P}\mathbf{m}\|_{L^2(\mathbb{R}^3)^3}$$

for arbitrary $v \in H^{1/2}(\Omega) \setminus \{0\}$. Taking the supremum over all $v \in H^{1/2}(\Omega) \setminus \{0\}$ we obtain

$$\|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\Omega)} \lesssim \|\mathcal{P}\mathbf{m}\|_{L^2(\mathbb{R}^3)^3}. \tag{35}$$

For $\mathcal{P}\widetilde{\mathbf{m}} = -\nabla\widetilde{p}$, the representation $\widetilde{p} = S(-\nabla \cdot \widetilde{\mathbf{m}})$ and the variational equality (34) for $\mathcal{P}\mathbf{m} = -\nabla p$ imply

$$(\mathcal{P}\mathbf{m}, \mathcal{P}\widetilde{\mathbf{m}})_{L^2(\mathbb{R}^3)^3} = -\langle \nabla \cdot \mathbf{m}, \widetilde{p} \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)} = -\langle \nabla \cdot \mathbf{m}, V(-\nabla \cdot \widetilde{\mathbf{m}}) \rangle_{\widetilde{H}^{-1/2}(\Omega) \times H^{1/2}(\Omega)}.$$

The choice $\widetilde{\mathbf{m}} = \mathbf{m}$ yields $(\mathcal{P}\mathbf{m}, \mathcal{P}\mathbf{m})_{L^2(\mathbb{R}^3)^3} = \|\nabla \cdot \mathbf{m}\|_V^2$, which allows us to conclude the proof. □

Note that as a direct consequence of Proposition 12, we may indeed rewrite our energy functional from (17)–(18) in the form (29)–(30). The semi-norm $\|\!\|\cdot\|\!\| \lesssim \|\cdot\|_{\mathcal{H}}$ of (30) is obviously continuous on $\mathcal{H}$. Since $\mathcal{H} \subseteq L^2(\Omega)^2$ is by construction a continuously embedded Hilbert space, the thin-film problem fits into the general setting of Section 1. It only remains to prove the assumptions (A1) and (A2), which guarantees the existence of minimizers and allows us to apply our convergent numerical scheme of Section 2.

Lemma 13. *For the micromagnetic thin-film model* (TF), *the energy functional of* $e(\mathbf{m})$ *of* (29)–(30) *and the penalized energy* $e_\varepsilon(\mathbf{m}) = e(\mathbf{m}) + \frac{1}{2\varepsilon}\|(|\mathbf{m}| - 1)_+\|_{L^2(\Omega)}^2$ *satisfy the coercivity assumptions* (A1) *and* (A2). *In particular,* (TF) *has a minimizer* $\mathbf{m}^*$.

Proof. For the energy $e(\mathbf{m})$ of the constrained problem, the coercivity follows by an easy argument: Let $(\mathbf{m}^n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ with $\lim_{n\to\infty} \|\mathbf{m}^n\|_{\mathcal{H}} = \infty$. Since $|\mathbf{m}^n| \leq 1$ almost everywhere in Ω, we have that

$$\|\mathbf{m}^n\|_{L^2(\Omega)} \leq |\Omega|^{1/2} \quad \text{and} \quad |(\mathbf{f}, \mathbf{m}^n)_{L^2(\Omega)}| \leq \|\mathbf{f}\|_{L^1(\Omega)}.$$

¿From the boundedness of the L^2-norm of $\mathbf{m}^n$, we get $\lim_{n\to\infty} \|\nabla \cdot \mathbf{m}^n\|_{\widetilde{H}^{-1/2}(\Omega)} = \infty$. From the boundedness of the scalar product $(\mathbf{f}, \mathbf{m}^n)_{L^2(\Omega)}$ and the equivalence of norms $\|\cdot\|_V \simeq \|\cdot\|_{\widetilde{H}^{-1/2}(\Omega)}$ we obtain

$$e(\mathbf{m}^n) \geq C\|\nabla \cdot \mathbf{m}^n\|_{\widetilde{H}^{-1/2}(\Omega)}^2 - \|\mathbf{f}\|_{L^1(\Omega)} \longrightarrow \infty.$$

The proof of assumption (A2) is a little bit more involved as also the L^2-norm of a sequence $(\mathbf{m}^n)_{n\in\mathbb{N}}$ with $\limsup_n \|\mathbf{m}^n\|_{\mathcal{H}} = \infty$ may be possibly unbounded. From equivalence of norms $\|\nabla \cdot \mathbf{m}\|_V \eqsim \|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\omega)}$, we get the existence of a constant $C_1 > 0$ such that

$$e_\varepsilon(\mathbf{m}) \geq C_1\|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\Omega)}^2 + \frac{q}{2}\|\mathbf{m}_2\|_{L^2(\Omega)}^2 + \frac{1}{2\varepsilon}\|(|\mathbf{m}| - 1)_+\|_{L^2(\Omega)}^2 - (\mathbf{f}, \mathbf{m})_{L^2(\Omega)}.$$

Let $\Omega_\geq$ denote the set where $|\mathbf{m}| \geq 1$, and $\Omega_<$ its complement, i.e. $|\mathbf{m}(x)| < 1$ for $x \in \Omega_<$. We use Hölder's inequality to estimate the linear contribution

$$\int_\Omega \mathbf{f} \cdot \mathbf{m} \, dx \leq \|\mathbf{f}\|_{L^2(\Omega)} \|\mathbf{m}\|_{L^2(\Omega_\geq)} + \|\mathbf{f}\|_{L^1(\Omega)}.$$

Next, for the penalty energy contribution it holds that

$$
\begin{aligned}
\int_\Omega (|\mathbf{m}| - 1)_+^2 \, dx &= \int_{\Omega_\geq} |\mathbf{m}|^2 - 2|\mathbf{m}| + 1 \, dx \\
&\geq \|\mathbf{m}\|_{L^2(\Omega_\geq)}^2 - 2\|\mathbf{m}\|_{L^2(\Omega_\geq)} |\Omega_\geq|^{1/2} + |\Omega_\geq| \\
&\geq \|\mathbf{m}\|_{L^2(\Omega_\geq)}^2 - 2\|\mathbf{m}\|_{L^2(\Omega_\geq)} |\Omega|^{1/2},
\end{aligned}
\tag{36}
$$

Applying these inequalities we obtain

$$
\begin{aligned}
e_\varepsilon(\mathbf{m}) &\geq \frac{1}{2\varepsilon} \|(|\mathbf{m}| - 1)_+\|_{L^2(\Omega)}^2 - (\mathbf{f}, \mathbf{m})_{L^2(\Omega)} + C_1 \|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\Omega)}^2 \\
&\geq \frac{1}{2\varepsilon} \left(\|\mathbf{m}\|_{L^2(\Omega_\geq)}^2 - 2|\Omega|^{1/2} \|\mathbf{m}\|_{L^2(\Omega_\geq)} \right) - \|\mathbf{f}\|_{L^2(\Omega)} \|\mathbf{m}\|_{L^2(\Omega_\geq)} - \|\mathbf{f}\|_{L^1(\Omega)} + C_1 \|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\Omega)}^2 \\
&= \frac{1}{2\varepsilon} \left(\|\mathbf{m}\|_{L^2(\Omega_\geq)}^2 - \left(2|\Omega|^{1/2} + 2\varepsilon\|\mathbf{f}\|_{L^2(\Omega)} \right) \|\mathbf{m}\|_{L^2(\Omega_\geq)} - 2\varepsilon\|f\|_{L^1(\Omega)} \right) + C_1 \|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\Omega)}^2.
\end{aligned}
$$

Defining the constants $C_2 = 2|\Omega|^{1/2} + 2\varepsilon\|\mathbf{f}\|_{L^2(\Omega)}$ and $C_3 = 2\varepsilon\|\mathbf{f}\|_{L^1(\Omega)}$, we conclude

$$e_\varepsilon(\mathbf{m}) \geq \frac{1}{2\varepsilon} \left(\|\mathbf{m}\|_{L^2(\Omega_\geq)}^2 - C_2 \|\mathbf{m}\|_{L^2(\Omega_\geq)} - C_3 \right) + C_1 \|\nabla \cdot \mathbf{m}\|_{\widetilde{H}^{-1/2}(\Omega)}^2. \tag{37}$$

From

$$\|\mathbf{m}\|_{L^2(\Omega)} \leq \|\mathbf{m}\|_{L^2(\Omega_\geq)} + \|\mathbf{m}\|_{L^2(\Omega_<)} \leq \|\mathbf{m}\|_{L^2(\Omega_\geq)} + |\Omega|^{1/2}, \tag{38}$$

we conclude the proof with the following observations: Let $(\mathbf{m}^n) \subseteq \mathcal{H}$ be a sequence of magnetizations with $\limsup_n \|\mathbf{m}^n\|_\mathcal{H} = \infty$. Then $\limsup_n \|\nabla \cdot \mathbf{m}^n\|_{\widetilde{H}^{-1/2}(\Omega)} = \infty$ or $\limsup_n \|\mathbf{m}^n\|_{L^2(\Omega)} = \infty$. Hence, at least one contribution, either the $\|\cdot\|_V$-norm of the divergence or the L^2-norm of $\mathbf{m}$, will cause the energy to be unbounded. $\qquad\square$

So far, we have seen that the thin-film minimization problem fits into the general setting of Sections 1 and 2. The existence of solutions $\mathbf{m}^*$ follows for any $q \geq 0$ by the direct method of calculus of variations. The following lemma gives additional information on the uniqueness of the solutions: If $q > 0$, then the energy semi-norm $\|\|\cdot\|\|$ is in fact positive definite and thus a norm. Uniqueness of the minimizer $\mathbf{m}^*$ follows in this case.

Lemma 14. *Let $\mathbf{m} \in \mathcal{H}$ with $\nabla \cdot \mathbf{m} = 0 \in \widetilde{H}^{-1/2}(\Omega)$ and $\mathbf{m}_2 = 0 \in L^2(\Omega)$. Then $\mathbf{m} = 0 \in \mathcal{H}$.*

Proof. Note that $\nabla \cdot \mathbf{m} = 0 \in L^2(\Omega)$. Therefore $\mathbf{m}$ is an element of the space $H^1(\nabla \cdot; \Omega)$. Furthermore the extension

$$\widehat{\mathbf{m}} = \begin{cases} \mathbf{m}(x) & x \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

is an element of $H^1(\nabla \cdot; \mathbb{R}^2)$ with $\nabla \cdot \widehat{\mathbf{m}} = 0$. Indeed, taking a sequence $(\mathbf{m}^n)_{n \in \mathbb{N}} \subset \mathcal{D}(\Omega) \subset \mathcal{H}$ with $\mathbf{m}^n \to \mathbf{m}$ in $\mathcal{H}$, we obtain for any test function $\varphi \in \mathcal{D}(\mathbb{R}^2)$

$$(\widehat{\mathbf{m}}, \nabla\varphi)_{L^2(\mathbb{R}^2)} = \lim_{n \to \infty} (\widehat{\mathbf{m}}^n, \nabla\varphi)_{L^2(\mathbb{R}^2)} = -\lim_{n \to \infty} (\nabla \cdot \mathbf{m}^n, \varphi)_{L^2(\Omega)} = (\nabla \cdot \mathbf{m}, \varphi)_{L^2(\Omega)} = 0.$$

Next, we show that $\widehat{\mathbf{m}} = 0 \in L^2(\mathbb{R}^2)$ which in turn implies $\mathbf{m} = 0 \in \mathcal{H}$. For this, we choose a sequence of mollifiers $\psi_\varepsilon \in \mathcal{D}(\mathbb{R}^2)$ with $\text{supp}(\psi) \subseteq [-\varepsilon, \varepsilon]^2$ and

$$\psi_\varepsilon \star \widehat{\mathbf{m}} \longrightarrow \widehat{\mathbf{m}} \in L^2(\mathbb{R}^2) \text{ as } \varepsilon \to 0 \tag{39}$$

13

componentwise, where $\star$ denotes the convolution of functions, cf. [20, Theorem 3.4]. ¿From e.g. [22, Lemma 2.13], we know

$$\psi_\varepsilon \star \widehat{\mathbf{m}} \in H^1(\nabla\cdot;\mathbb{R}^2) \cap \mathcal{D}(\mathbb{R}^2)^2 \quad \text{with} \quad \nabla \cdot (\psi_\varepsilon \star \widehat{\mathbf{m}}) = \psi_\varepsilon \star (\nabla \cdot \widehat{\mathbf{m}}) = 0.$$

¿From $\widehat{\mathbf{m}}_2 = 0 \in L^2(\mathbb{R}^2)$, we additionally know

$$\psi_\varepsilon \star \widehat{\mathbf{m}} = \begin{pmatrix} \psi_\varepsilon \star \widehat{\mathbf{m}}_1 \\ 0 \end{pmatrix}$$

which trivially implies $\frac{\partial(\psi_\varepsilon \star \widehat{\mathbf{m}}_2)}{\partial x_2} = 0$. Together with $\nabla \cdot (\psi_\varepsilon \star \widehat{\mathbf{m}}) = 0$ this means

$$\frac{\partial(\psi_\varepsilon \star \widehat{\mathbf{m}}_1)}{\partial x_1} = 0.$$

Since $(\psi_\varepsilon \star \widehat{\mathbf{m}}_1) \in \mathcal{D}(\mathbb{R}^2)$ is a smooth function with compact support, $\nabla(\psi_\varepsilon \star \widehat{\mathbf{m}}_1) = 0$ already implies $\psi_\varepsilon \star \widehat{\mathbf{m}}_1 = 0$. $\qquad\square$

Remark. *For $q > 0$, the norm $\|\cdot\|_{\mathcal{H}}$ is obviously stronger than $\|\cdot\|$ and these two norms are not equivalent. To see this, let $\Omega = (-1,1)^2$ and $X, Y \in \mathcal{D}(\mathbb{R})$ with $\operatorname{supp} X \subseteq (-1,1)$ and $\operatorname{supp} Y \subseteq (-1,1)$. Assume $X \neq 0$ and $Y \neq 0$. Define the sequence $(\mathbf{m}^n)_{n\in\mathbb{N}}$ by*

$$\mathbf{m}_1^n(x_1, x_2) := -nX(x_1)Y'(nx_2), \qquad \mathbf{m}_2^n(x_1, x_2) := X'(x_1)Y(nx_2).$$

Then

$$\nabla \cdot \mathbf{m}^n = 0, \quad \mathbf{m}^n \in \mathcal{D}(\Omega), \quad \|\mathbf{m}^n\|^2 = \|\mathbf{m}_2^n\|_{L^2(\Omega)}^2 = \frac{1}{n}\|X'\|_{L^2(\mathbb{R})}\|Y\|_{L^2(\mathbb{R})} \tag{40}$$

and

$$\|\mathbf{m}^n\|_{\mathcal{H}}^2 = \|\mathbf{m}^n\|_{L^2(\Omega)}^2 = \frac{1}{n}\|X'\|_{L^2(\mathbb{R})}\|Y\|_{L^2(\mathbb{R})} + n\|X\|_{L^2(\mathbb{R})}\|Y'\|_{L^2(\mathbb{R})}. \tag{41}$$

This shows that a bound of the form $\|\mathbf{m}\|_{\mathcal{H}} \leq C\|\mathbf{m}\|$ cannot hold on $\mathcal{H}$. In fact, the example shows that norm equivalence cannot hold for any *domain $\Omega \subset \mathbb{R}^2$. Moreover, since $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ is continuously and bijectively embedded into $(\mathcal{H}, \|\cdot\|)$, Banach's open mapping theorem predicts that $(\mathcal{H}, \|\cdot\|)$ is not complete.* $\qquad\square$

So far, we have proved (i) and (ii) of Theorem 11. As a last issue in this section, we show continuous dependence of the solutions on the given data. To that end we need the corresponding variational inequality. The following standard result can be found e.g. in [17] and is formulated here in our setting for the convenience of the reader.

Lemma 15. *Let H be a Hilbert space with continuous semi-scalar product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$. Furthermore, let $\Phi \in L(H;\mathbb{R})$ and let $\mathcal{A} \subseteq H$ denote a closed and convex subset. Given the energy functional*

$$e(u) = \frac{1}{2}\langle\!\langle u, u \rangle\!\rangle - \Phi(u),$$

an element $u^* \in \mathcal{A}$ is a minimizer, i.e.

$$e(u^*) \leq e(v) \quad \text{for all } v \in \mathcal{A}, \tag{42}$$

if and only if u^* satisfies the variational inequality

$$\langle\!\langle u^*, u^* - v \rangle\!\rangle \leq \Phi(u^* - v) \quad \text{for all } v \in \mathcal{A}. \quad \square \tag{43}$$

Lemma 16. *If $\mathbf{m}^1$ and $\mathbf{m}^2$ are solutions to (TF) for applied fields $\mathbf{f}^1$ and $\mathbf{f}^2$, respectively, then*

$$\|\mathbf{m}^1 - \mathbf{m}^2\| \leq \sqrt{2}\|\mathbf{f}^1 - \mathbf{f}^2\|_{L^1(\Omega)}^{1/2}. \tag{44}$$

Proof. To obtain the estimate (44), note that solving the minimization problem (M) is equivalent to solving the variational inequality

$$\langle\!\langle \mathbf{m}\,,\,\mathbf{m}-\mathbf{w}\rangle\!\rangle \leq (\mathbf{f},\mathbf{m}-\mathbf{w}) \quad \forall \mathbf{w} \in \mathcal{A},$$

see Lemma 15. Since $\mathbf{m}^1$ solves the variational inequality for $\mathbf{f}^1$ and $\mathbf{m}^2$ solves the variational inequality for $\mathbf{f}^2$, we have

$$\begin{aligned}
\|\mathbf{m}^1 - \mathbf{m}^2\|^2 &= \langle\!\langle \mathbf{m}^1\,,\,\mathbf{m}^1-\mathbf{m}^2\rangle\!\rangle + \langle\!\langle \mathbf{m}^2\,,\,\mathbf{m}^2-\mathbf{m}^1\rangle\!\rangle \leq (\mathbf{f}^1,\mathbf{m}^1-\mathbf{m}^2)_{L^2} + (\mathbf{f}^2,\mathbf{m}^2-\mathbf{m}^1)_{L^2}\\
&= (\mathbf{f}^1-\mathbf{f}^2,\mathbf{m}^1-\mathbf{m}^2)_{L^2}\\
&\leq \|\mathbf{f}^1-\mathbf{f}^2\|_{L^1(\Omega)}\|\mathbf{m}^1-\mathbf{m}^2\|_{L^\infty(\Omega)}.
\end{aligned}$$

The observation $|\mathbf{m}^1(x)-\mathbf{m}^2(x)| \leq 2$ and taking the square root concludes the proof. $\qquad\square$

4 Discretization and experimental analysis

The last step to apply our numerical scheme is to provide a conforming discretization of the energy space $\mathcal{H}$. Note that $H^1(\nabla\cdot;\Omega) \subseteq \mathcal{H}$ dense. The Raviart-Thomas finite elements introduced in [23] to discretize $H^1(\nabla\cdot;\Omega)$ are thus a natural choice for the discretization of such an energy space, cf. [13].

4.1. The space of Raviart-Thomas finite elements. Let $\mathcal{T}_h$ be a regular triangulation of the domain Ω in the sense of Ciarlet, i.e.:

- Each element $T_j \in \mathcal{T}_h$ is a non-degenerate and closed triangle,
- $\mathcal{T}_h$ covers Ω, i.e. $\overline{\Omega} = \bigcup_{T \in \mathcal{T}_h} T$,
- The intersection $T_i \cap T_j$, for $i \neq j$, is either empty, a common vertex, or a common edge.

The global mesh size h is defined by $h = \max_{T \in \mathcal{T}_h} \text{diam}(T)$. Moreover, the set of all edges of a triangulation is denoted by $\mathcal{E}_h$ and $\mathcal{E}_h^\Omega$ is the set of all interior edges.

We define the space of lowest order Raviart-Thomas finite elements by

$$RT^0(\mathcal{T}_h) = \{\mathbf{m}_h \in \mathcal{P}^1(\mathcal{T}_h)^2 \,|\, [\mathbf{m}_h \cdot \mathbf{n}_E]_E = 0 \,\forall E \in \mathcal{E}_h^\Omega \text{ and } \mathbf{m}_h \cdot \mathbf{n} = 0 \text{ on } \Gamma\},$$

where $\mathcal{P}^1(\mathcal{T}_h)$ denotes the space of piecewise linear and discontinuous functions, $\mathbf{n}_E$ denotes a normal vector on the edge E, and $[\cdot]_E$ denotes the jump across an edge of the triangulation.

We stress that the crucial property $[\mathbf{m}_h \cdot \mathbf{n}_E]_E = 0$ ensures the $H^1(\nabla\cdot;\Omega)$ conformity of the discrete space $RT^0(\mathcal{T}_h)$. Since $H^1(\nabla\cdot;\Omega) \subseteq \mathcal{H}$, the set

$$\mathcal{A}_h := \{\mathbf{m}_h \in RT^0(\mathcal{T}_h) \,|\, |\mathbf{m}_h| \leq 1 \text{ a.e.}\}$$

is a conforming discretization of our admissible set $\mathcal{A}$.

Next, we describe the standard basis of $RT^0(\mathcal{T}_h)$. Each interior edge $E \in \mathcal{E}_h^\Omega$ belongs to precisely two elements T_+ and T_-. For such an edge E, let P_+ and P_- be the vertices of T_+ and T_- opposite E; in other words, $T_\pm = \text{conv}(E \cup \{P_\pm\})$ as shown in Figure 1. For each $E \in \mathcal{E}_h^\Omega$, we define

$$\psi_E = \begin{cases} \pm\frac{|E|}{2|T_\pm|}(x - P_\pm), & \text{for } x \in T_\pm \\ 0, & \text{elsewhere} \end{cases} \tag{45}$$

and notice that the jump $[\psi_E \cdot \mathbf{n}]$ across any edge vanishes. This implies $\psi_E \in RT^0(\mathcal{T}_h)$. Moreover, it can be shown that the set

$$\mathcal{B} = \{\psi_E \,|\, E \in \mathcal{E}_h^\Omega\}$$

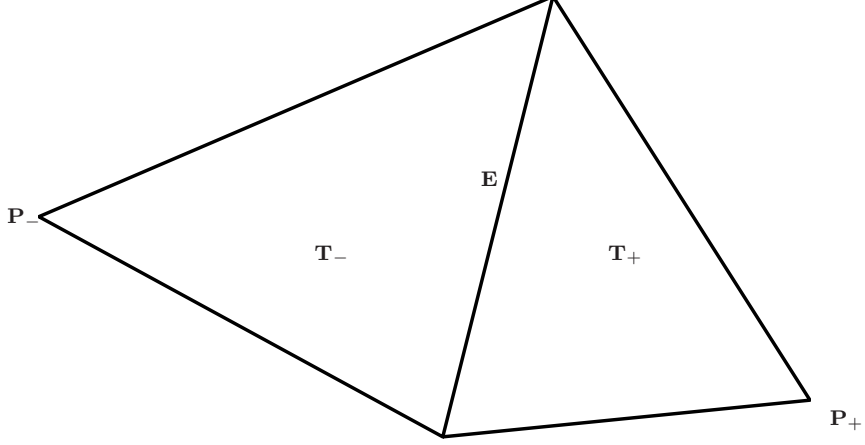is a basis of $RT^0(\mathcal{T}_h)$, cf. [13].

Figure 1: Each interior edge E belongs to precisely two triangles T_+ and T_-. The points opposite of E are denoted by P_+ and P_-, respectively.

Remark. *The variational inequality* (43) *for both, the continuous and the discrete constrained minimization problem, provides a tool to derive a priori error estimates. Indeed,* (43) *applied on the continuous and the discrete level gives us for arbitrary* $\mathbf{w}_h \in \mathcal{A}_h$

$$
\begin{aligned}
\|\mathbf{m}^* - \mathbf{m}_h^0\|^2 &= \langle\!\langle \mathbf{m}^*, \mathbf{m}^* \rangle\!\rangle - 2\langle\!\langle \mathbf{m}^*, \mathbf{m}_h^0 \rangle\!\rangle + \langle\!\langle \mathbf{m}_h^0, \mathbf{m}_h^0 \rangle\!\rangle \\
&\leq \langle\!\langle \mathbf{m}^*, \mathbf{m}_h^0 \rangle\!\rangle + (f, \mathbf{m}^* - \mathbf{m}_h^0)_{L^2} - 2\langle\!\langle \mathbf{m}^*, \mathbf{m}_h^0 \rangle\!\rangle + \langle\!\langle \mathbf{m}_h^0, \mathbf{w}_h^0 \rangle\!\rangle + (f, \mathbf{m}_h^0 - \mathbf{w}_h^0)_{L^2} \\
&= -\langle\!\langle \mathbf{m}^*, \mathbf{m}_h^0 \rangle\!\rangle + (f, \mathbf{m}^* - \mathbf{w}_h)_{L^2} + \langle\!\langle \mathbf{m}_h^0, \mathbf{w}_h \rangle\!\rangle \\
&= \langle\!\langle \mathbf{m}_h^0, \mathbf{w}_h - \mathbf{m}^* \rangle\!\rangle + (f, \mathbf{m}^* - \mathbf{w}_h)_{L^2} \\
&= \langle\!\langle \mathbf{m}^* - \mathbf{m}_h^0, \mathbf{m}^* - \mathbf{w}_h \rangle\!\rangle - \langle\!\langle \mathbf{m}^*, \mathbf{m}^* - \mathbf{w}_h \rangle\!\rangle + (f, \mathbf{m}^* - \mathbf{w}_h)_{L^2}.
\end{aligned}
$$

Applying the trivial inequality $2ab \leq a^2 + b^2$ *to the term* $\langle\!\langle \mathbf{m}^* - \mathbf{m}_h^0, \mathbf{m}^* - \mathbf{w}_h \rangle\!\rangle$, *we see*

$$
\|\mathbf{m}^* - \mathbf{m}_h^0\|^2 \leq \frac{1}{2}\|\mathbf{m}^* - \mathbf{m}_h^0\|^2 + \frac{1}{2}\|\mathbf{m}^* - \mathbf{w}_h\|^2 - \langle\!\langle \mathbf{m}^*, \mathbf{m}_h^* - \mathbf{w}_h \rangle\!\rangle + (f, \mathbf{m}^* - \mathbf{w}_h)_{L^2}
$$

$$
\|\mathbf{m}^* - \mathbf{m}_h^0\|^2 \leq \|\mathbf{m}^* - \mathbf{w}_h\|^2 - 2\langle\!\langle \mathbf{m}^*, \mathbf{m}^* - \mathbf{w}_h \rangle\!\rangle + 2(f, \mathbf{m}^* - \mathbf{w}_h)_{L^2}.
$$

Since $\mathbf{w}_h \in \mathcal{A}_h$ *was arbitrary the last inequality still holds when taking the infimum over all* $\mathbf{w}_h \in \mathcal{A}_h$ *on the right-hand side. Choosing* $\mathbf{w}_h = \Pi_h \mathbf{m}^*$ *with a suitable interpolation operator* $\Pi_h : H^1(\Omega) \to RT^0(\mathcal{T}_h)$, *one can prove an a priori rate of*

$$
\|\mathbf{m}^* - \mathbf{m}_h^0\| = \mathcal{O}(\sqrt{h}),
$$

provided that $\mathbf{m}^* \in H^1(\Omega)^2$ *with* $\nabla \cdot \mathbf{m} \in H^1(\Omega)$. *Note, however, that norm convergence of the constrained problem does not imply norm convergence of the penalized system.* □

4.2. A simple damped Newton algorithm. The penalized energy functional $e_\varepsilon(\mathbf{m})$ is Fréchet differentiable with derivative

$$
De_\varepsilon(\mathbf{m})(\mathbf{w}) = (\nabla \cdot \mathbf{m}, \nabla \cdot \mathbf{w})_V + q(\mathbf{m}_2, \mathbf{w}_2) - (\mathbf{f}, \mathbf{w})_{L^2(\Omega)} + \left(\frac{(|\mathbf{m}| - 1)_+}{\varepsilon |\mathbf{m}|} \mathbf{m}, \mathbf{w} \right)_{L^2(\Omega)}
$$

(see [11] for the detailed derivation of the last term). Since $e_\varepsilon(\mathbf{m})$ is convex, we have that finding a global minimizer of (M_h^ε) is equivalent to solving the variational problem

$$
De(\mathbf{m}_h^\varepsilon)(\mathbf{w}_h) = 0 \quad \forall \mathbf{w}_h \in RT^0(\mathcal{T}_h).
$$

We apply Newton's algorithm to solve this equation. Note that this derivative of the penalized energy functional is not continuously differentiable due to the penalty energy. Hence no classical result

on convergence of Newton's algorithm applies. Let $\mathcal{T}_h$ denote some regular triangulation with $\#\mathcal{E}_h^\Omega =: N_D$ interior edges. Given some coefficient vector $\mathbf{x} \in \mathbb{R}^{N_D}$ the discrete Euler-Lagrange equation reads

$$(\nabla \cdot \sum_{i=1}^{N_D} \mathbf{x}_i \psi_i, \nabla \cdot \mathbf{w}_h)_V + q(\sum_{i=1}^{N_D} \mathbf{x}_i \psi_{i,2}, \mathbf{w}_{h,2})_{L^2} - (f, \mathbf{w}_h)_{L^2} + \frac{1}{\varepsilon}(\lambda_h^\varepsilon \sum_{i=1}^{N_D} \mathbf{x}_i \psi_i, \mathbf{w}_h) = 0$$

for all $\mathbf{w}_h \in RT^0(\mathcal{T}_h)$. Obviously this equation holds for all $\mathbf{w}_h \in RT^0(\mathcal{T}_h)$ if and only if it holds for all basis functions $(\psi_i)_{i=1}^{N_D}$ of $RT^0(\mathcal{T}_h)$. Hence we seek the zeros of the discrete function $F : \mathbb{R}^{N_D} \to \mathbb{R}^{N_D}$ defined by

$$F_j(\mathbf{x}) = (\nabla \cdot \mathbf{m}_h^\varepsilon(\mathbf{x}), \nabla \cdot \psi_j)_V + q(\mathbf{m}_{h,2}^\varepsilon(\mathbf{x}), \psi_{j,2})_{L^2} + \frac{1}{\varepsilon}(\lambda_h^\varepsilon \mathbf{m}_h^\varepsilon(\mathbf{x}), \psi_j)_{L^2} - (f, \psi_j)_{L^2}. \qquad (46)$$

The notation $\mathbf{m}_h^\varepsilon(\mathbf{x})$ indicates that the discrete magnetization $\mathbf{m}_h^\varepsilon = \mathbf{m}_h^\varepsilon(\mathbf{x})$ depends on the given coefficient vector $\mathbf{m}_h^\varepsilon = \sum_{i=1}^{N_D} \mathbf{x}_i \psi_i$. A crucial step is the computation of the derivative of the function F. In equation (46), the derivative of the first two scalar products can be computed easily. The fourth contribution vanishes, since it is just a constant. The third term, however, is more involved. The derivative of $\frac{(|\mathbf{m}|-1)_+}{|\mathbf{m}|}$ is not defined classically at the points where $|\mathbf{m}_h^\varepsilon(\mathbf{x})| = 1$. However, since the scalar product is computed by use of numerical quadrature, this exceptional situation is not expected to be encountered numerically at any quadrature point. In our implementation, points where $|\mathbf{m}(x)| = 1$ are treated in the same way as points where $|\mathbf{m}(x)| < 1$. The following formula for the Jacobian of the non-linear contribution is obtained by straight forward calculations.

Lemma 17. *The Jacobian* $\mathbf{DF_{NL}} \in \mathbb{R}^{N_D \times N_D}$ *of the function* $F_{NL} : \mathbb{R}^{N_D} \to \mathbb{R}^{N_D}$ *defined by*

$$\mathbf{x} \mapsto (g(\mathbf{x}), \psi_j)_{L^2} \quad \text{with} \quad g(\mathbf{x}) := \frac{(|\mathbf{m}_h^\varepsilon(\mathbf{x})| - 1)_+}{|\mathbf{m}_h^\varepsilon(\mathbf{x})|} \mathbf{m}_h^\varepsilon(\mathbf{x})$$

is given by

$$(\mathbf{DF_{NL}})(\mathbf{x})_{i,j} = (\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}_i}, \psi_j)_{L^2},$$

where at any $x \in \Omega$ *with* $|\mathbf{m}_h^\varepsilon(\mathbf{x})(x)| \neq 1$ *the derivative of* g *either reads* $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}_i}(x) = 0$ *in the case* $|\mathbf{m}_h^\varepsilon(\mathbf{x})(x)| < 1$ *or, if* $|\mathbf{m}_h^\varepsilon(\mathbf{x})(x)| > 1$,

$$\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}_i}(x) = \psi_i - \begin{pmatrix} \frac{1}{|\mathbf{m}_h^\varepsilon|} - \frac{\mathbf{m}_{h,1}^\varepsilon}{|\mathbf{m}_h^\varepsilon|^3} & -\frac{\mathbf{m}_{h,1}^\varepsilon \mathbf{m}_{h,2}^\varepsilon}{|\mathbf{m}_h^\varepsilon|^3} \\ -\frac{\mathbf{m}_{h,1}^\varepsilon \mathbf{m}_{h,2}^\varepsilon}{|\mathbf{m}_h^\varepsilon|^3} & \frac{1}{|\mathbf{m}_h^\varepsilon|} - \frac{\mathbf{m}_{h,2}^\varepsilon}{|\mathbf{m}_h^\varepsilon|^3} \end{pmatrix} \cdot \psi_i. \qquad \square$$

In our simulation runs, we found that the Newton algorithm did not converge in all cases. For smooth functions the following modified algorithm, often referred to as relaxed Newton algorithm or damped Newton method, converges globally. However, this is at the cost of decreased order of convergence.

Damped Newton algorithm:

Let $\mathbf{x}^{(0)} \in \mathbb{R}^{N_D}$ denote some initial value and set $\ell = 0$.

(i) Evaluate $F(\mathbf{x}^{(\ell)})$ and compute the derivative $DF^{(\ell)}$.

(ii) Compute the search direction $\delta \in \mathbb{R}^{N_D}$ by solving the linear system $DF^{(\ell)}\delta = -F(\mathbf{x}^{(\ell)})$.

(iii) Find minimal k such that $|F(\mathbf{x}^{(\ell)} + 0.5^k \delta)| < |F(\mathbf{x}^{(\ell)})|$

(iv) Define $\mathbf{x}^{(\ell+1)} := \mathbf{x}^{(\ell)} + 0.5^k \delta$

(v) Either stop or $\ell \mapsto \ell + 1$ and goto (i)

Output: An approximation $\mathbf{x}$ to some root of the function F.

Note that step (iii) ensures a reduction of the residual in each step. As a stopping criterion we simply check $|F(\mathbf{x})| < 10^{-8}$. More sophisticated stopping criteria that, e.g., also deal with the possibility of not
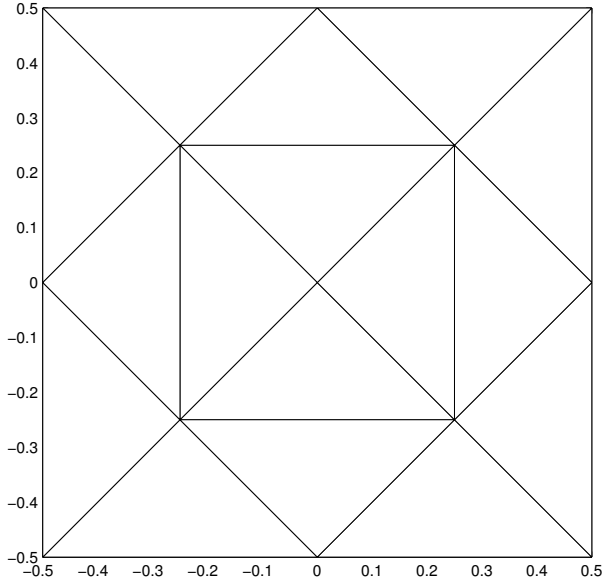
Figure 2: Initial mesh $\mathcal{T}_h$ of the simulation domain $\Omega = (-0.5, 0.5)^2$.

being able to find a root up to rounding errors could be chosen. This, however, was not necessary in any of our simulation runs so that we could work with the simpler criterion.

Let $\mathcal{T}_h$ be some given mesh with triangles $T_1, \ldots, T_{N_T}$ and $N_D = \#\mathcal{E}_h^\Omega$ interior edges. Let χ_{T_j} denote the characteristic function of a triangle T_j. Note that $\{\chi_{T_1}, \ldots, \chi_{T_{N_T}}\}$ is a basis of the space $\mathcal{P}^0(\mathcal{T}_h)$ of piecewise constant functions on the mesh. Then, we define the system matrices

$$
\begin{aligned}
\mathbf{V} &\in \mathbb{R}^{N_T \times N_T}, & \mathbf{V}_{i,j} &:= (\chi_{T_j}, V\chi_{T_i})_{L^2}, \\
\mathbf{Q} &\in \mathbb{R}^{N_D \times N_D}, & \mathbf{Q}_{i,j} &:= (\psi_{j,2}, \psi_{i,2})_{L^2}, \\
\mathbf{D} &\in \mathbb{R}^{N_T \times N_D}, & \mathbf{D}_{i,j} &:= (\nabla \cdot \psi_j)|_{T_i}.
\end{aligned}
\tag{47}
$$

We define the vectors $\mathbf{b}_j := (\mathbf{f}, \psi_j)_{L^2}$ and $\mathbf{a}_j(\mathbf{x}) := \left(\frac{(|\mathbf{m}_h^\varepsilon(\mathbf{x})|-1)_+}{|\mathbf{m}_h^\varepsilon(\mathbf{x})|} |\mathbf{m}_h^\varepsilon(\mathbf{x})|, \psi_j \right)_{L^2}$. Finally the evaluation of the function $F(\mathbf{x})$ and the derivative matrix $\mathbf{DF}(\mathbf{x})$ read

$$
F(\mathbf{x}) = \mathbf{D}^T \mathbf{V} \mathbf{D} \mathbf{x} + q\mathbf{Q}\mathbf{x} + \mathbf{a}(\mathbf{x}) - \mathbf{b},
$$
$$
\mathbf{DF}(\mathbf{x}) = \mathbf{D}^T \mathbf{V} \mathbf{D} + q\mathbf{Q} + \mathbf{DF_{NL}}(\mathbf{x}).
$$

Note that the matrix $\mathbf{V}$ is dense. We apply matrix compression techniques well-known from boundary element analysis to store and work with the matrix in a data-sparse manner. Among the many strategies available, we use here hierachical matrices [14], which have log-linear complexity for storing $\mathbf{V}$ and realizing the matrix-vector multiplication $\mathbf{z} \mapsto \mathbf{V}\mathbf{z}$. Our implementation utilizes the HLib library (http://www.hlib.org) for the computation of the matrix $\mathbf{V}$.

4.3. Numerical experiments. We study the behavior of our algorithm with a simple set of experiments. We choose the sample to be the unit square $\Omega = (-0.5, 0.5)^2$. The applied field is constant—a standard assumption in thin-film micromagnetics [7]. Since for this non-linear problem no analytical solution is available, we estimate the error by comparing with a reference solution $\mathbf{m}_{ref}$ that is computed on a relatively fine grid. Throughout, the anisotropy parameter is $q = 1$.

4.3.1 Convergence in h

In this first experiment we are interested in the convergence as $h \to 0$ for fixed ε. To be precise, we choose $\varepsilon = 10^{-2}$ and we compute the solutions $\mathbf{m}_h^\varepsilon$ on a sequence of uniformly refined meshes of Ω. The applied field is $\mathbf{f} = (1, -0.5)^T$. The initial mesh with $\#\mathcal{T}_h = 16$ triangles is shown in Figure 2.
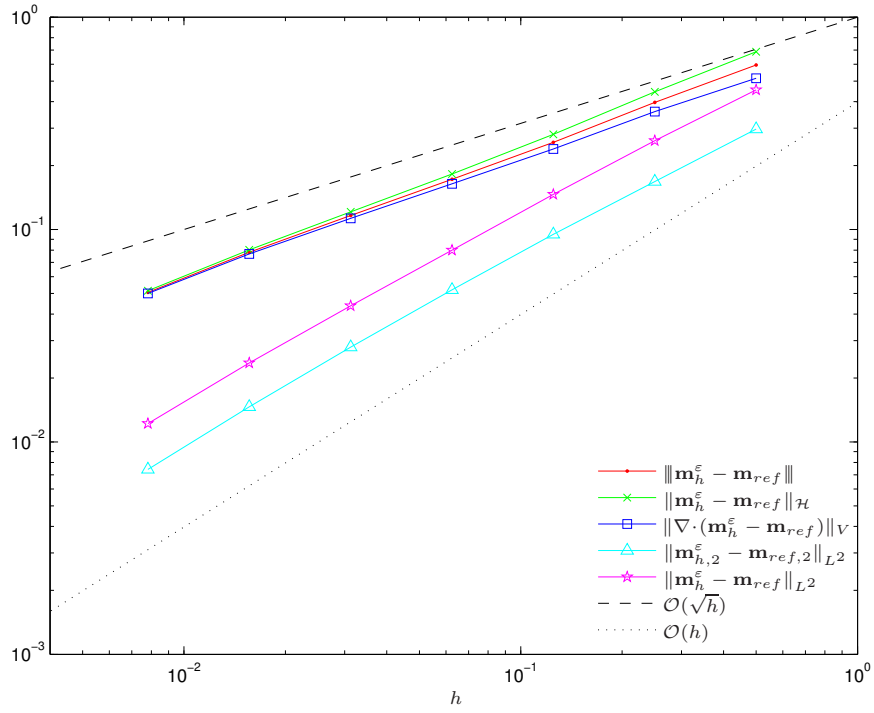
Figure 3: Error $\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}$ measured in various norms and plotted versus the mesh size h. The slopes corresponding to $\mathcal{O}(h)$ and $\mathcal{O}(\sqrt{h})$ in the double logarithmic scale are shown for reference.
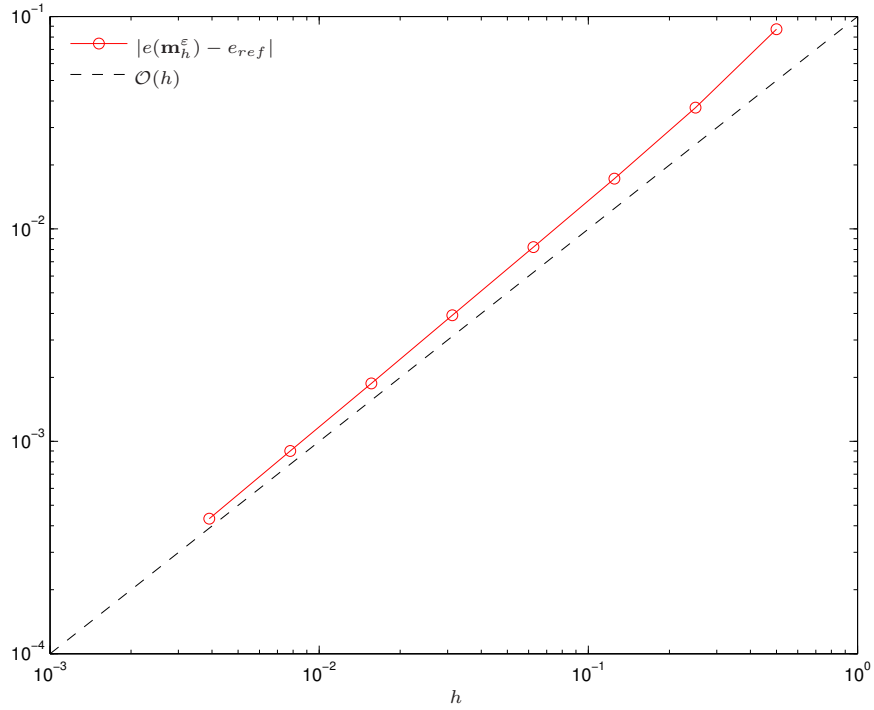

Figure 4: Error in the energy $|e(\mathbf{m}_h^\varepsilon) - e_{ref}| \approx |e(\mathbf{m}_h^\varepsilon) - e(\mathbf{m}_0^\varepsilon)|$ versus the mesh size h. The slope corresponding to $\mathcal{O}(h)$ in the double logarithmic scale is shown for reference.
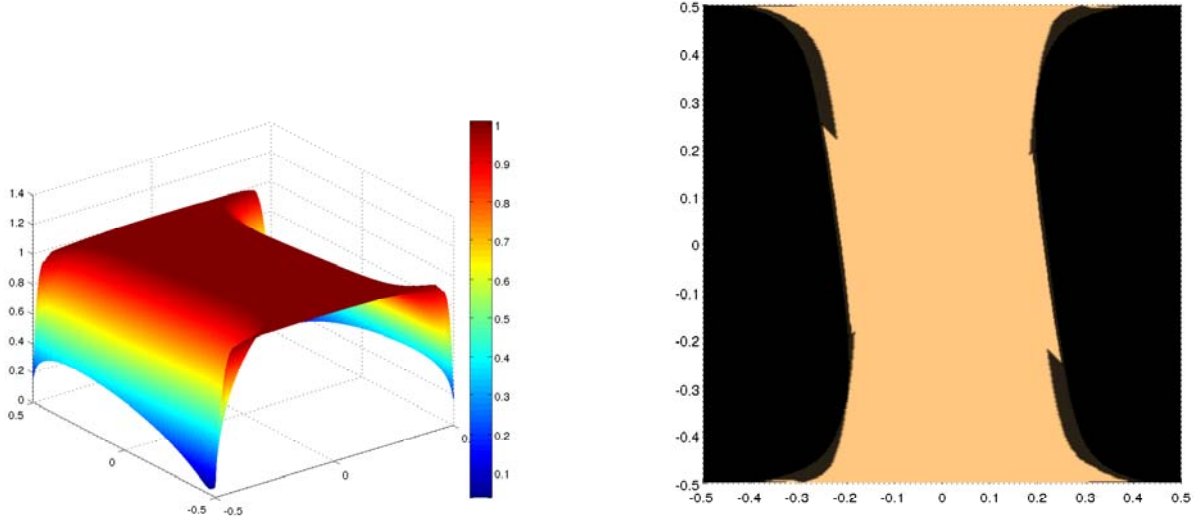
Figure 5: Left: Length of the magnetization $\mathbf{m}_{ref}$. Right: Region of Ω where the penalty term is active, i.e. $|\mathbf{m}_{ref}| > 1$ somewhere in the corresponding triangle.

The reference solution $\mathbf{m}_{ref}$ was computed on a mesh $\mathcal{T}_h$ with $262,144$ triangles, which corresponds to $392,704$ degrees of freedom for the discrete space $RT^0(\mathcal{T}_h)$. Figure 5 shows the length of the discrete magnetization and the region of Ω where the penalty term is active.

Figure 3 shows the estimated error in the energy norm $\||\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}\|| \approx \||\mathbf{m}_h^\varepsilon - \mathbf{m}_0^\varepsilon\||$ as well as the full space norm $\|\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}\|_{\mathcal{H}}$. We also plot the contributions of the L^2-norms $\|\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}\|_{L^2(\Omega)}$ and $\|\mathbf{m}_{h,2}^\varepsilon - \mathbf{m}_{ref,2}\|_{L^2(\Omega)}$ as well as the error of the divergence in the V-norm $\|\nabla \cdot \mathbf{m}_h^\varepsilon - \nabla \cdot \mathbf{m}_{ref}\|_V$. All quantities are plotted versus the mesh size h. Finally, Figure 4 shows the energy error $|e(\mathbf{m}_h^\varepsilon) - e_{ref}|$. We computed the reference quantity $e_{ref} = -0.5815508709$ by extrapolation from the sequence of energies $e(\mathbf{m}_h^\varepsilon)$ with Aitken's Δ^2-method.

The dominant error contribution is the error in the divergence $\|\nabla \cdot (\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref})\|_V$. We observe an asymptotic behavior of $\mathcal{O}(\sqrt{h})$. The error in the L^2-norm decays linearly $\mathcal{O}(h)$ and is of higher order. Also the error in the first component, which is not controlled by the energy functional, decays at the same linear rate. The error in the energy decays linearly at $\mathcal{O}(h)$—which is what we expect since the energy is a quadratic quantity.

4.3.2 Convergence in ε

In this second experiment we study the convergence of $\mathbf{m}_h^\varepsilon$ to $\mathbf{m}_h^0$ as $\varepsilon \to 0$. We compute discrete and penalized solutions to the same problem as in the first experiment in Section 4.3.1. But this time we compute the solutions $\mathbf{m}_h^\varepsilon$ for varying penalty parameter on a fixed mesh with $\#\mathcal{T}_h = 4,096$ triangles. This corresponds to $6,080$ degrees of freedom of the discrete space $RT^0(\mathcal{T}_h)$ and a mesh size of $h = 0.0312$. The reference solution $\mathbf{m}_{ref}$ was computed with a value of $\varepsilon = 3.05 \cdot 10^{-5}$.

Figure 6 shows the estimated error in the energy norm $\||\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}\|| \approx \||\mathbf{m}_h^\varepsilon - \mathbf{m}_0^\varepsilon\||$ as well as the full space norm $\|\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}\|_{\mathcal{H}}$. We also plot the contributions of the L^2-norms $\|\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}\|_{L^2(\Omega)}$ and $\|\mathbf{m}_{h,2}^\varepsilon - \mathbf{m}_{ref,2}\|_{L^2(\Omega)}$ as well as the error of the divergence in the V-norm $\|\nabla \cdot \mathbf{m}_h^\varepsilon - \nabla \cdot \mathbf{m}_{ref}\|_V$. All quantities are plotted versus the penalty parameter ε. Figure 7 shows the error $|e(\mathbf{m}_h^\varepsilon) - e_{ref}|$ in the energy. We computed the reference quantity $e_{ref} = -0.5760956532$ by extrapolation from the sequence of energies $e(\mathbf{m}_h^\varepsilon)$ with Aitken's Δ^2-method.

We observe two major differences to the error with respect to h-refinements. First, we observe that the error in the energy norm decays with order $1/2$, i.e. $\||\mathbf{m}_h^\varepsilon - \mathbf{m}_h^0\|| = \mathcal{O}(\sqrt{\varepsilon})$. This is similar to the error decay with respect to h. In contrast, however, the error in the energy also only decays at the same order $1/2$ with respect to ε. This is a bit surprising since the energy is a quadratic quantity. Apparently, the error from the approximation by the penalty method is reflected in the energy mainly through the linear L^2-contribution $(\mathbf{f}, \mathbf{m})_{L^2}$. The second difference is that the error in the L^2-norm and the error of
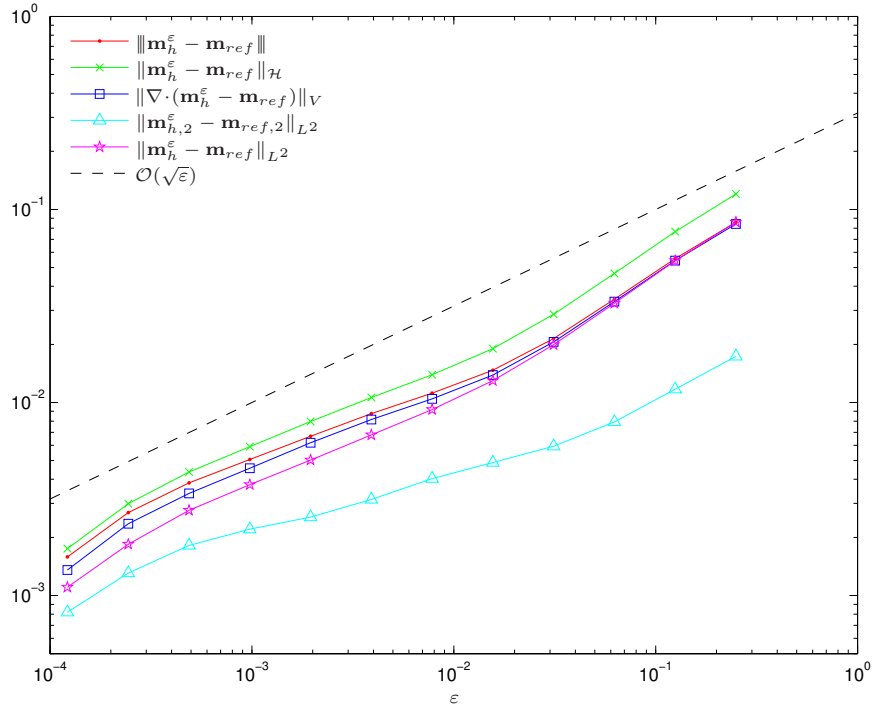
Figure 6: Error $\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}$ measured in various norms and plotted versus the penalty parameter ε. The slope corresponding to $\mathcal{O}(\sqrt{\varepsilon})$ in the double logarithmic scale is shown for reference.
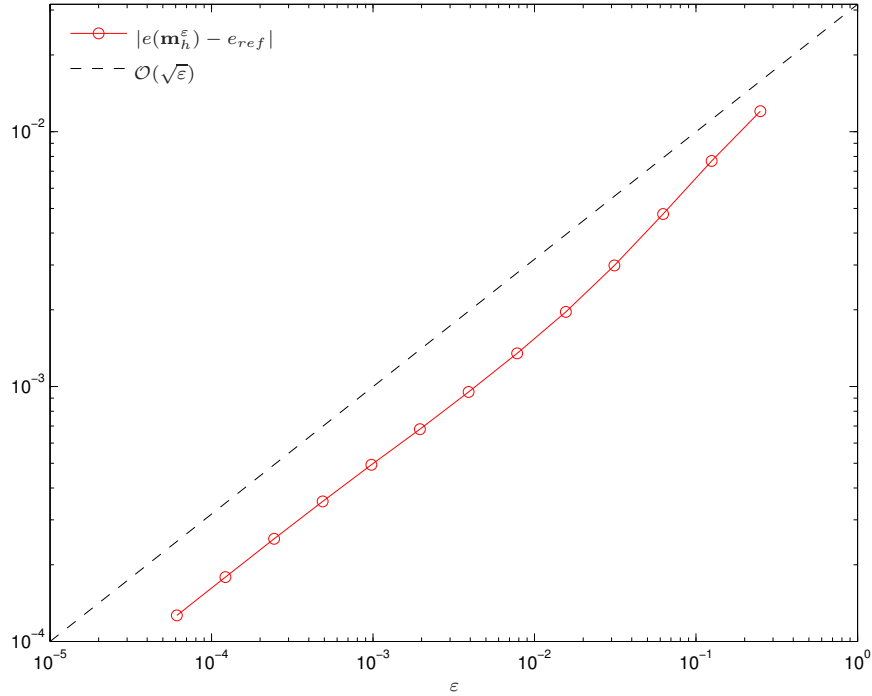

Figure 7: Energy error $|e(\mathbf{m}_h^\varepsilon) - e_{ref}| \approx |e(\mathbf{m}_h^\varepsilon) - e(\mathbf{m}_h^0)|$ versus the penalty parameter ε. The slope corresponding to $\mathcal{O}(\sqrt{\varepsilon})$ in the double logarithmic scale is shown for reference.
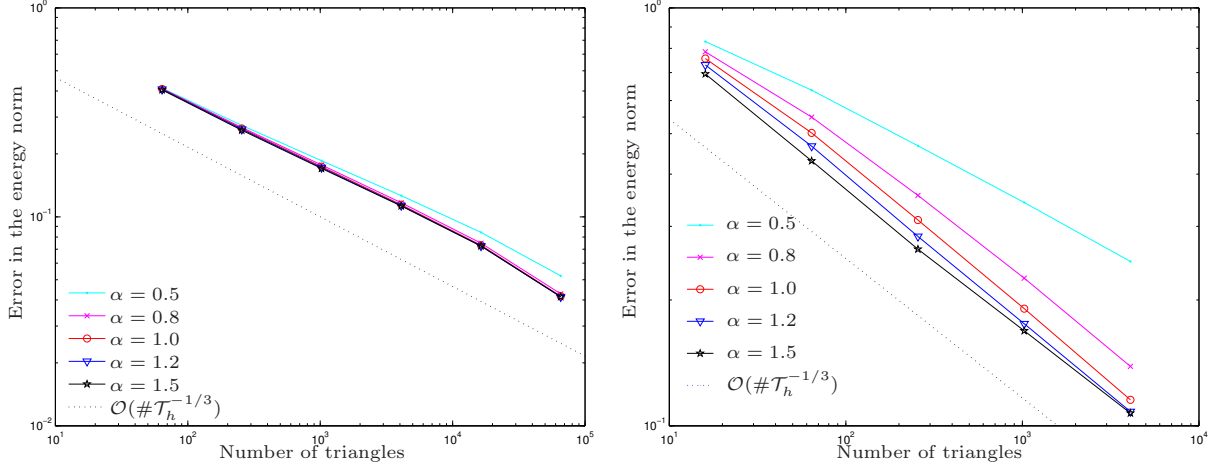
Figure 8: Error in the energy norm $\|\mathbf{m}_h^\varepsilon - \mathbf{m}_{ref}\| \approx \|\mathbf{m}_h^\varepsilon - \mathbf{m}^*\|$ versus the number of triangles $\#\mathcal{T}_h$ for different choices of $\varepsilon = h^\alpha$. The slope corresponding to $\mathcal{O}(\#\mathcal{T}_h^{1/3})$ in the double logarithmic scale is shown for reference. Left: Applied field $\mathbf{f} = (1, -0.5)^T$. Right: Applied field $\mathbf{f} = (2, 2)^T$.

the divergence in the V-norm are of the same order.

4.3.3 Choice of $\varepsilon = h^\alpha$

It is an open question how to balance the parameters ε and h in numerical calculations. So far, we only looked at the error contributions independently. Putting everything together might yield a different picture. To empirically understand the optimal choice of $\varepsilon = h^\alpha$, we perform a series of calculations for varying $\alpha \in \{0.5, 0.8, 1.0, 1.2, 1.5\}$. We use the same geometry and parameters as in the first experiment in Section 4.3.1. The reference solution $\mathbf{m}_{ref}$ was computed on a mesh with $\#\mathcal{T}_h = 65,536$ triangles which corresponds to $98,048$ degrees of freedom. The penalty parameter of the reference solution is $\varepsilon = 6.52 \cdot 10^{-4}$.

Figure 8 (left) shows the error in the energy norm for the sequences of solutions $\mathbf{m}_h^\varepsilon$ with $\varepsilon = h^\alpha$ plotted versus the number of triangles $\#\mathcal{T}_h$ of the corresponding mesh. We clearly see that $\alpha = 0.5$ leads to a reduced order of convergence compared to the other choices. The plot is not clearly visible beyond $\alpha = 0.8$. This is because the discretization error dominates the total error. In order to obtain more conclusive data, in a second experiment, we applied a larger field $\mathbf{f} = (2, 2)^T$. This choice emphasizes certain effects. The error contributions from the discretization and the penalty scheme are well balanced for our purposes. Figure 8 (right) shows the results of this second simulation run. We observe that the choice of $\alpha = 1$ is optimal. Choice of $\alpha < 1$ leads to a reduced order of convergence; the choice of $\alpha > 1$ leads to the same order of convergence as $\alpha = 1$. We note that the energy norm error, in both simulation, is approximately $\mathcal{O}(\#\mathcal{T}_h^{-1/3})$. Recalling the relation $h \sim \#\mathcal{T}_h^{-1/2}$, this means that the convergence behavior $\mathcal{O}(\#\mathcal{T}_h^{-1/3})$ is of higher order than $\sqrt{h}$. One possible explanation is that, as the mesh size decreases, the active set is better resolved, see also the more detailed experimental analysis in [11]. We are not sure, however, whether this accelerated convergence effect is only a pre-asymptotic phenomenon due to an inadequate resolution of the active set at the beginning of the calculations.

5 Conclusions

In the first part of the paper we presented a general convergence result for penalty methods. The proof applies to a large class of quadratic minimization problems. We stress that it does not involve any optimality conditions, and the assumptions on the inequality constraints are very weak.

In the second part of the paper we analyzed the thin-film model in micromagnetics from [7]. In contrast to the original works [8, 9], our perspective is that of the numerical analyst. We construct an appropriate Hilbert space and prove well-posedness of the problem in this setting. We note that the

uniqueness of the solution in the case $q \neq 0$ is not stated in the prior works [7, 8, 9, 10]. The present work generalizes and extends the results of [11] and [12].

Parts of the dissertation [10] are also concerned with the numerical simulation of the same thin-film model problem. There, the author uses an interior point method to compute admissible approximations to the magnetization. While admissibility is mandatory in many applications, we feel that a good qualitative and even quantitative understanding of the present thin-film model can be obtained with numerical methods that relax the admissibility constraint. Additionally, we were able to give a mathematical convergence result for the algorithms presented here.

The numerical experiments show that there is still some work to do. The observation that uniform meshes lead to a sub-linear order of convergence is consistent with the observations in [10]. If the applied field is sufficiently weak so that the constraint $|\mathbf{m}| \leq 1$ is not active, the minimization problem reduces to a certain linear integral equation. ¿From the literature on boundary element methods it is then clear that the divergence $\nabla \cdot \mathbf{m}$ has generic singularities along the boundary of Ω. A heuristic adaptive algorithm developed in [11] shows that the rate of convergence can be improved. There are some open questions in this context, however, since adaptive mesh refinements towards the boundary of Ω is suboptimal for the approximation of $\mathbf{m}$ in the L^2-norm. Also a posteriori error estimators that measure the error with respect to h and ε are not justified rigorously.

Despite these open questions, the numerical experiments presented here did support our analytical results—convergence in the energy norm holds for arbitrary choices of $(h, \varepsilon) \to (0, 0)$. The system matrix $\mathbf{V}$ of (47) is the bottle-neck in the calculations. From that point of view, our simple Newton method is quite successful. The size of the computations is mainly limited by memory consumption and the building time of $\mathbf{V}$, i.e. by effects emerging from the nature of the problem.

References

[1] R. Adams, *Sobolev spaces*, Pure and Applied Mathematics, Vol. 65. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[2] C. Carstensen, A. Prohl, *Numerical analysis of relaxed micromagnetics by penalised finite elements*, Numer. Math., 90(1):65–99, 2001.

[3] C. Carstensen, D. Praetorius, *Numerical analysis for a macroscopic model in micromagnetics*, SIAM J. Numer. Anal., 42(6):2633–2651, 2005.

[4] C. Carstensen, D. Praetorius, *Effective simulation of a macroscopic model for stationary micromagnetics*, Comput. Methods Appl. Mech. Engrg., 194(2-5), 531–548, 2005.

[5] B. Dacorogna, *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*, Springer-Verlag, Berlin, 1989.

[6] J. Deny, J. L. Lions *Les espaces du type de Beppo Levi*, Ann. Inst. Fourier, Grenoble, 5:305–370, 1954.

[7] A. DeSimone, R. Kohn, S. Müller, F. Otto, R. Schäfer, *Two-dimensional modelling of soft ferromagnetic films*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci., 457(2016):2983–2991, 2001.

[8] A. DeSimone, R. Kohn, S. Müller, F. Otto, *A reduced theory for thin-film micromagnetics*, Comm. Pure Appl. Math., 55(11):1408–1460, 2002.

[9] A. DeSimone, R. Kohn, S. Müller, F. Otto, *Recent analytical developments in micromagnetics*, in Giorgo Bertotti and Isaak D. Mayergoyz, editors, *The science of hysteresis. Vol. II*, pages 269–381, Elsevier/Academic Press, Amsterdam, 2006.

[10] J. Drwenski, *Numerical Methods for a Reduced Model in Thin-Film Micromagnetics*, Dissertation, Rheinische Friedrichs-Wilhelms-Universität Bonn, 2008.

[11] S. Ferraz-Leite, *Quadratic minimization with non-local operators and non-linear constraints*, Dissertation, Vienna University of Technology, 2011.

[12] S. Ferraz-Leite, J.M. Melenk, D. Praetorius, *Reduced Model in Thin-Film Micromagnetics*, Proceedings MATHMOD 09 Vienna, I. Troch, F. Breitenecker (edts.), Argesim / Asim, ARGESIM Report no. 35, 2009.

[13] V. Girault, P. Raviart, *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics*, Springer-Verlag, Berlin, 1986.

[14] W. Hackbusch, *A sparse matrix arithmetic based on H-matrices. Part I: Introduction to H-matrices*, Computing, 62:89-108, 1999.

[15] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*, Springer, New York, 2009.

[16] K. Ito, K. Kunisch, *Lagrange Multiplier approach to variational problems and applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[17] D. Kinderlehrer, G. Stampacchia, *An introduction to variational inequalities and their applications*, volume 88 of *Pure and Applied Mathematics*, Academic Press Inc., [Harcourt Brace Jovanovich Publishers], New York, 1980.

[18] L. D. Landau, E. M. Lifschits, *On the theory of the dispersion of magnetic permeability in ferromagnetic bodies*, Phys. Zeitsch. der Sow., 8:153–169, 1935.

[19] M. Maischak, E. P. Stephan, *A priori error estimates for hp penalty BEM for contact problems in elasticity*, Comput. Methods Appl. Mech. Engrg. 196, no. 37-40, 3871–3880, 2007.

[20] W. McLean *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, 2000.

[21] J. Nocedal, S. Wright, *Numerical optimization*, Springer Series in Operations Research, Springer-Verlag, New York, 1999.

[22] D. Praetorius, *Analysis, Numerik und Simulation eines relaxierten Modellproblems zum Mikromagnetismus*, Dissertation, Vienna University of Technology, 2003.

[23] P.-A. Raviart, J. M. Thomas, *Primal hybrid finite element methods for 2nd order elliptic equations*, Math. Comp., 31(138):391–413, 1977.

[24] S. Sauter, C. Schwab *Boundary Element Methods*, Springer Verlag, Berlin, 2011.

[25] E. P. Stephan, *Boundary integral equations for screen problems in $\mathbb{R}^3$*, Integral Equations Operator Theory, 10(2):236–257, 1987