Comput Mech (2007) 39:663–672
DOI 10.1007/s00466-006-0105-2

ORIGINAL PAPER

On stabilized models in micromagnetics

Carsten Carstensen · Dirk Praetorius

Received: 17 December 2005 / Accepted: 25 March 2006 / Published online: 15 August 2006

Abstract The effective behaviour of stationary micromagnetic phenomena is modelled by a convexified Landau–Lifshitz minimization problem for the limit of large and soft magnets Ω with no exchange energy. The numerical simulation of the resulting minimization problem has to overcome difficulties caused by the pointwise side-constraint $|\mathbf{m}| \leq 1$ and the stray-field energy on the unbounded domain $\mathbb{R}^d$. A penalty method models the side-constraint and the exterior Maxwell equation is recast via a nonlocal integral operator $\mathcal{L}$. This paper gives an overview of the available results and implementational details.

Keywords Micromagnetics · Relaxation · Effective model · Adaptive algorithm · Newton-Raphson scheme

1 Introduction

Stationary micromagnetic phenomena of static or quasi-static processes are usually based on a variational model named after Landau and Lifshitz [4,17,19,20]. The magnetic body Ω is a bounded Lipschitz domain in $\mathbb{R}^d$ for $d = 2,3$ on which the microscopic vector-valued magnetization

$$\mathbf{m}_\alpha : \Omega \to \mathbb{R}^d$$

minimizes the magnetic energy $E_\alpha(\mathbf{m})$ subject to the constraint $|\mathbf{m}| =$ constant depending on the temperature. The effective magnetization vector $\mathbf{m} : \Omega \to \mathbb{R}^d$ is a spatial average of the microscopic magnetization $\mathbf{m}_\alpha$ and so averages out the small oscillations which $\mathbf{m}_\alpha$ is enforced to develop for small values of the exchange energy parameter $\alpha \geq 0$, cf. [18,21,22]. The energy comprises of four terms known as exchange energy, anisotropic energy, exterior field (or Zeeman) energy and stray-field (or magnetostatic) energy, namely

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

For large and soft magnets, the parameter α vanishes in the magnetic energy E_α. This is justified in [12], where it is proven that the effective model for $\alpha \to 0$ is $E_0(\mathbf{m})$ and the effective magnetization $\mathbf{m} : \Omega \to \mathbb{R}^d$ obeys the averaged constraint $|\mathbf{m}| \leq 1$ (with the aforementioned constant normalized to 1). In fact, E_0 is non-convex owing to the non-convex side restriction and infimizing sequences are enforced to develop oscillations of smaller and smaller length-scale (in contrast to E_α where $\alpha > 0$ determines some smallest length-scale). In the limit, the oscillations describe a measure rather than a Lebesgue function. The expected value of this limiting Young measure is the effective or macroscopic magnetization $\mathbf{m}$ in the model and can in fact directly be computed by some relaxed energy density.

C. Carstensen (✉)
Department of Mathematics,
Humboldt-Universität
zu Berlin, Unter den Linden 6,
10099 Berlin, Germany
e-mail: cc@math.hu-berlin.de

D. Praetorius
Vienna University of Technology, Institute for Analysis
and Scientific Computing, Wiedner Hauptstraße 8-10,
1040 Vienna, Austria
e-mail: Dirk.Praetorius@tuwien.ac.at

Note that this effective model is the mathematical foundation of the so-called *phase theory* in micromagnetics [17]. This paper follows [8,9,11] and adopts the effective model directly. The relaxed problem (RP) [12,29] reads: minimize

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

subject to the side-constraint

$$|\mathbf{m}(x)| \le 1 \quad \text{for almost every } x \in \Omega. \tag{3}$$

Given some direction $\mathbf{e} \in \mathbb{R}^d$, called easy axis with $|\mathbf{e}| = 1$, and an orthonormal basis $(\mathbf{e}, \mathbf{z}_2, \mathbf{z}_3, \ldots, \mathbf{z}_d)$ of $\mathbb{R}^d$, for uniaxial materials such as Cobalt the anisotropic energy density ϕ^{**} in (2) reads

$$\phi^{**}(\mathbf{m}) = \frac{1}{2}\sum_{j=2}^{d}(\mathbf{m}\cdot\mathbf{z}_j)^2 \quad \text{for all } \mathbf{m} \in \mathbb{R}^d. \tag{4}$$

This specifies the first term out of three of (2), and we restrict to the uniaxial case in the following. The second is a linear relation with a given applied magnetization $\mathbf{f} \in L^2(\Omega;\mathbb{R}^d)$; here and below we employ standard notation for Lebesgue spaces: hence, $\mathbf{f} \in L^2(\Omega)$ e.g. means that $\mathbf{f}$ is measurable and L^2 integrable (i.e. $\int_\Omega |\mathbf{f}(x)|^2\,\mathrm{d}x < \infty$), while $L^2(\Omega;\mathbb{R}^d)$ denotes $L^2(\Omega)\times\cdots\times L^2(\Omega)$ in d components. The third term in (2) models the stray field energy in $\mathbb{R}^d$. Given a magnetization $\mathbf{m}$ (as the argument in $E_0^{**}(\mathbf{m})$), the associated magnetic potential u solves a Laplace equation

$$\Delta u = \operatorname{div}\mathbf{m} \quad \text{in } \mathbb{R}^d \text{ in the sense of distributions.} \tag{5}$$

Note that (5) involves interface conditions $[\partial u/\partial\mathbf{n}] = -\mathbf{m}\cdot\mathbf{n}$ for the jump on $\partial\Omega$. Here and below, $\mathbf{m}$ is an L^∞ function on Ω and extended by zero outside the magnetic domain Ω. Then, ∇u is uniquely determined and belongs to $L^2(\mathbb{R}^d;\mathbb{R}^d)$; hence the third energy term is finite. This concludes the short description of the effective model via (RP). In fact, (RP) has solutions [12] which are unique [8,25].

For $d = 2$, the numerical analysis of the effective model was initiated by [11]. Therein, the effective magnetization $\mathbf{m}$ is discretized by a $\mathcal{T}$-piecewise constant ansatz and the side-constraint is enforced by a penalty method, where $\mathcal{T}$ is a triangulation of Ω. The entire space $\mathbb{R}^2$ in (5) is replaced by a bounded domain $\widehat{\Omega}$ containing Ω, and the potential equation is solved with respect to homogeneous boundary values in $H_0^1(\widehat{\Omega})$. For the solution of the potential equation, [11] uses nonconforming $P1$ elements, since the coupling of conforming $P0$–$P1$ finite elements leads to instabilities. These can be overcome by adding a certain stabilization term based on jumps over edges/faces of the triangulation $\mathcal{T}$, cf. [13]. While [11] provides a priori and a posteriori error control, [13] states only some a priori error estimates. In all these results the L^2-norm of the error $\boldsymbol{\delta}_h = \mathbf{m} - \mathbf{m}_h$ is only controlled in the $\mathbf{z}_j$-directions, i.e. there are *no* estimates for $\|\boldsymbol{\delta}_h\cdot\mathbf{e}\|_{L^2(\Omega)}$. The only estimate concerning the latter term can be found in [28]. Therein a (much more complicated) higher order discretization of $\mathbf{m}$ and $u \in H_0^1(\widehat{\Omega})$ is considered, which is stabilized again with certain jump terms. The a priori result states a reduced convergence order $\|\boldsymbol{\delta}_h\cdot\mathbf{e}\|_{L^2(\Omega)} = \mathcal{O}(h^{1/2})$, whereas $\|\boldsymbol{\delta}_h\cdot\mathbf{z}_j\|_{L^2(\Omega)} = \mathcal{O}(h)$ as in [11,13]. Unfortunately, the technical details of the proof remain unclear, and no numerical experiments for this discretization are given.

In contrast to [11,13,28], the entire space $\mathbb{R}^d$ in (5) is considered in [8,9] by use of an integral operator $\mathcal{L}$. For a $\mathcal{T}$-piecewise constant magnetization $\mathbf{m}_h$, the potential $u_h = \mathcal{L}\mathbf{m}_h$ is computed exactly. The main advantage of this ansatz is that one only has to deal with *one* discretization for $\mathbf{m}_h$ instead of with a coupled discretization as before. Moreover, it follows that the a priori and a posteriori results from [11] essentially carry over to this discretization and can be generalized to cover the $3D$ case, as well [8]. As before, the error estimates are only concerned with $\|\boldsymbol{\delta}_h\cdot\mathbf{z}_j\|_{L^2(\Omega)}$. Recently, it could be proven that, for $d = 2$, this discrete model leads to *fully weak* convergence $\boldsymbol{\delta}_h \rightharpoonup 0$ in $L^2(\Omega;\mathbb{R}^d)$ and, moreover, the use of an appropriate stabilization yields $\|\boldsymbol{\delta}_h\|_{L^2(\Omega;\mathbb{R}^2)} = \mathcal{O}(h^{1/2})$, cf. [10].

This paper reviews the full model and its discretization from [8,9] accompanied by an a priori and a posteriori error analysis in Sect. 2. The focus is then on numerical aspects of the solution of the discrete model. We compare the performance of the simple Newton–Raphson scheme with a gradient method for which global convergence is proven in Sect. 3. Adaptive mesh-refinement is discussed and underlined with a numerical example from [8]. Section 4 summarizes a few observations and comments on future developments.

2 Preliminaries and error estimates

This section presents the effective model proposed for the numerical simulation in more detail. It introduces the reformulation of the stray field energy contribution by a non-local integral operator $\mathcal{P}$ as well as the penalty formulation for the constraint $|\mathbf{m}| \le 1$. Recalling the discrete model from [8–10], we continue with the review of the corresponding a priori and a posteriori error control.

2.1 The exact model

Let $\mathbf{m} = (m_1, \dots, m_d)$ be a magnetization, and denote by $G : \mathbb{R}^d \backslash \{0\} \to \mathbb{R}$ the Newtonian kernel,

$$G(x) := \begin{cases} \gamma_2^{-1} \log |x| & \text{for } d = 2, \\ \gamma_d^{-1}/(2-d)\,|x|^{2-d} & \text{for } d > 2, \end{cases} \tag{6}$$

for $x \neq 0$, where the positive constant γ_d is the surface measure of the unit sphere, e.g. $\gamma_2 = 2\pi$, $\gamma_3 = 4\pi$. Given the convolution operator

$$\mathcal{L}\mathbf{m} := \sum_{j=1}^{d} (\partial G/\partial x_j) * m_j \quad \text{in } \mathbb{R}^d,$$

one can prove that $u = \mathcal{L}\mathbf{m} \in H^1_{\ell oc}(\mathbb{R}^d)$ satisfies $\mathcal{P}\mathbf{m} := \nabla u \in L^2(\mathbb{R}^d;\mathbb{R}^d)$ and (5). The Helmholtz projection operator

$$\mathcal{P} : L^2(\mathbb{R}^d;\mathbb{R}^d) \to L^2(\mathbb{R}^d;\mathbb{R}^d)$$

is the L^2 orthogonal projection onto the linear and closed subspace of all the gradients [14,22,27]. Hence, the stray field energy in $E_0^{**}(\mathbf{m})$ reads

$$\int_{\mathbb{R}^d} |\nabla u|^2 \,\mathrm{d}x = \int_{\mathbb{R}^d} (\mathcal{P}\mathbf{m}) \cdot \mathbf{m} \,\mathrm{d}x.$$

The Gâteaux derivative (also called first variation) of E_0^{**} yields the Problem (*RP*): Find $(\lambda, \mathbf{m}) \in L^2(\Omega) \times L^2(\Omega;\mathbb{R}^d)$ such that

$$\mathcal{P}\mathbf{m} + \nabla\phi^{**}(\mathbf{m}) + \lambda\mathbf{m} = \mathbf{f} \quad \text{a.e. in } \Omega, \tag{7}$$

$$\lambda \geq 0,\ |\mathbf{m}| \leq 1,\ \lambda(1 - |\mathbf{m}|) = 0 \quad \text{a.e. in } \Omega. \tag{8}$$

The variable λ is the Lagrange multiplier for the side-constraint $|\mathbf{m}| \leq 1$, and (8) are the associated Kuhn–Tucker conditions. Note that (7) has to be solved on the *bounded* magnetic domain Ω only, although $\mathcal{P}\mathbf{m}$ has global support $\mathbb{R}^d$. This problem (*RP*) is well-posed in the sense that there exists a unique solution $\mathbf{m} \in L^2(\Omega;\mathbb{R}^d)$. According to convexity, $\mathbf{m}$ is the unique minimizer of E_0^{**} satisfying the side-constraint (3).

2.2 The discrete model

The spatial discretization of (RP) is done with a penalized Galerkin method which is based on a finite partition $\mathcal{T}$ of Ω into measurable subsets (of Ω). We suppose that, for $T \in \mathcal{T}$, the interior of T is a connected Lipschitz domain. Then, $\mathcal{L}^0(\mathcal{T})$ denotes the linear subspace of $\mathcal{T}$-piecewise constants; one such function is the mesh-size function $h := h_{\mathcal{T}} \in \mathcal{L}^0(\mathcal{T})$ defined by

$$h|_T := h_T := \operatorname{diam}(T) \text{ for all } T \in \mathcal{T}.$$

[Recall that $(\cdot)|_T$ denotes the restriction of a function $(\cdot)$ onto T.] Another example would be $\mathbf{f}_{\mathcal{T}} \in \mathcal{L}^0(\mathcal{T};\mathbb{R}^d)$, the piecewise integral mean of the given right-hand side $\mathbf{f} \in L^2(\Omega;\mathbb{R}^d)$ which is defined by

$$\mathbf{f}_{\mathcal{T}}|_T := |T|^{-1} \int_T \mathbf{f}(x)\,\mathrm{d}x \quad \text{for all } T \in \mathcal{T}. \tag{9}$$

Suppose that $\sigma(\cdot,\cdot)$ is a positive semi-definite bilinear form on $\mathcal{L}^0(\mathcal{T})^d$. Then, the (stabilized) discrete problem ($RP_{\varepsilon,h}$) reads as follows: Given a penalization parameter $\varepsilon \in \mathcal{L}^0(\mathcal{T})$ with $\varepsilon > 0$, find $\mathbf{m}_h \in \mathcal{L}^0(\mathcal{T})^d$ such that

$$\begin{aligned} &\langle \mathcal{P}\mathbf{m}_h + \nabla\phi^{**}(\mathbf{m}_h) + \lambda_h\mathbf{m}_h \,;\, \mathbf{n}_h\rangle_{L^2} + \sigma(\mathbf{m}_h, \mathbf{n}_h) \\ &\quad = \langle \mathbf{f} \,;\, \mathbf{n}_h\rangle_{L^2} \end{aligned} \tag{10}$$

for all $\mathbf{n}_h \in \mathcal{L}^0(\mathcal{T})^d$, where the discrete Lagrange multiplier $\lambda_h \in \mathcal{L}^0(\mathcal{T})$ is defined by

$$\lambda_h = \varepsilon^{-1} \frac{(|\mathbf{m}_h| - 1)_+}{|\mathbf{m}_h|} \quad \text{with } (\cdot)_+ := \max\{\cdot, 0\}. \tag{11}$$

The discrete model ($\mathrm{RP}_{\varepsilon,h}$) can be reformulated as a minimization problem to apply the direct method of the calculus of variations and to show well-posedness, i.e. the unique existence of solutions [8]. The numerical computation of the unique solution $\mathbf{m}_h$ is the topic of the subsequent sections.

Remark 1 (a) Note that, according to (11), ($\mathrm{RP}_{\varepsilon,h}$) leads to the solution of a nonlinear system of equations. In [8,9] this is done via a Newton–Raphson method which performs very efficiently, although convergence could not be proven.

(b) In Sect. 3 we will consider the stabilization

$$\sigma(\mathbf{m}_h, \mathbf{n}_h) = \langle h\mathbf{m}_h \,;\, h\mathbf{n}_h\rangle_{L^2} \tag{12}$$

to prove global convergence of a gradient method.

(c) The consideration of functions $\varepsilon, h \in \mathcal{L}^0(\mathcal{T})$ instead of scalars reflects the treatment of adaptive mesh-refinement to obtain $\mathcal{T}$.

Remark 2 Let $\mathcal{T}$ be a regular triangulation of $\Omega \subset \mathbb{R}^2$ in the sense of Ciarlet. Suppose that the elements $T \in \mathcal{T}$ are triangles, and denote by $\mathcal{S}^1(\mathcal{T})$ the usual $P1$ finite element space consisting of all continuous and $\mathcal{T}$-piecewise affine functions. In [10], the stabilization

$$\begin{aligned} \sigma(\mathbf{m}_h, \mathbf{n}_h) = {} & \tfrac{1}{2}\langle (id - \Pi)\mathbf{m}_h \,;\, (id - \Pi)\mathbf{n}_h\rangle_{L^2} \\ & + \frac{1}{2C}\langle h\nabla\Pi\mathbf{m}_h \,;\, h\nabla\Pi\mathbf{n}_h\rangle_{L^2} \end{aligned} \tag{13}$$

is used to prove that $\|\mathbf{m} - \mathbf{m}_h\|_{L^2(\Omega)} = \mathcal{O}(h^{1/2})$ provided $\mathbf{m}, \lambda\mathbf{m} \in H^1(\Omega;\mathbb{R}^2)$. Here, Π denotes the L^2 projection

onto $\mathcal{S}^1(\mathcal{T})^d$, and $C > 0$ is an appropriate constant which only depends on the shape of the elements in $\mathcal{T}$, but neither on size nor number. [In fact, $C > 0$ is chosen with respect to an inverse estimate.] This result is the first full convergence result for the lowest order discretization. However, we have to stress that we are unaware of any regularity properties of $\mathbf{m}$ and λ.

2.3 A priori error control

Throughout this section, let $(\lambda, \mathbf{m})$ and $(\lambda_h, \mathbf{m}_h)$ denote solutions of (RP) and $(RP_{\varepsilon,h})$, respectively. Then, there is a $\mathcal{T}$- and σ-independent constant $C > 0$ such that

$$\begin{aligned}
&\|\mathcal{P}\mathbf{m} - \mathcal{P}\mathbf{m}_h\|^2_{L^2(\mathbb{R}^d)} + \|\nabla\phi^{**}(\mathbf{m}) - \nabla\phi^{**}(\mathbf{m}_h)\|^2_{L^2(\Omega)} \\
&\quad + \|\lambda\mathbf{m} - \lambda_h\mathbf{m}_h\|^2_{L^2(\Omega)} \\
&\quad + \sigma(\mathbf{m}_\mathcal{T} - \mathbf{m}_h, \mathbf{m}_\mathcal{T} - \mathbf{m}_h) + \sigma(\mathbf{m}_h, \mathbf{m}_h) \\
&\le C(1 + \|\varepsilon\|_{L^\infty(\Omega)})\Big(\|\mathbf{m} - \mathbf{m}_\mathcal{T}\|^2_{L^2(\Omega)} \\
&\quad + \|\nabla\phi^{**}(\mathbf{m}) - (\nabla\phi^{**}(\mathbf{m}))_\mathcal{T}\|^2_{L^2(\Omega)} \\
&\quad + \|\lambda\mathbf{m} - (\lambda\mathbf{m})_\mathcal{T}\|^2_{L^2(\Omega)} + \sigma(\mathbf{m}_\mathcal{T}, \mathbf{m}_\mathcal{T})\Big) \\
&\quad + C\Big(\|\varepsilon\|_{L^\infty(\Omega)}\|\sqrt{\varepsilon}\,\lambda\mathbf{m}\|^2_{L^2(\Omega)} + \sigma((\lambda\mathbf{m})_\mathcal{T}, (\lambda\mathbf{m})_\mathcal{T})\Big) \\
&\quad + \sigma(\mathbf{m}_h, (\lambda\mathbf{m})_\mathcal{T} - \lambda_h\mathbf{m}_h). \qquad (14)
\end{aligned}$$

Recall that here and below $(\cdot)_\mathcal{T}$ denotes the $\mathcal{T}$-piecewise integral mean (9). Hence the first three terms on the right-hand side of (14) are best approximation errors. The last term $C\|\varepsilon\|_{L^\infty(\Omega)}\|\sqrt{\varepsilon}\lambda\mathbf{m}\|^2_{L^2(\Omega)}$ of order $\mathcal{O}(\varepsilon^2)$ is the penalty error. Furthermore, the estimate involves some stabilization errors. The most problematic term is the last stabilization term on the right-hand side of (14), since it is *not* an a priori term. For the proposed stabilization terms in (12) and (13), $\sigma(\mathbf{m}_h, (\lambda\mathbf{m})_\mathcal{T} - \lambda_h\mathbf{m}_h)$ can be absorbed on the left-hand side: for (12), there holds

$$\begin{aligned}
&\sigma(\mathbf{m}_h, (\lambda\mathbf{m})_\mathcal{T} - \lambda_h\mathbf{m}_h) \\
&\quad \le \frac{1}{2}\sigma(\mathbf{m}_h, \mathbf{m}_h) + \frac{1}{2}\sigma((\lambda\mathbf{m})_\mathcal{T}, (\lambda\mathbf{m})_\mathcal{T}) - \sigma(\mathbf{m}_h, \ell_h)
\end{aligned}$$

and $\sigma(\mathbf{m}_h, \ell_h) \ge 0$. For (13), it can be proven that

$$\sigma(\mathbf{m}_h, \mathbf{m}_h) \le \frac{1}{2}\sigma(\mathbf{m}_h, \mathbf{m}_h) + \frac{1}{2}\|\ell - \ell_h\|_2^2 + \frac{1}{2}\|\ell - \ell_\mathcal{T}\|_2^2.$$

If we suppose $\mathbf{m}$ and λ to be smooth, i.e. $\mathbf{m}, \lambda\mathbf{m} \in H^1(\Omega;\mathbb{R}^d)$, then (14) verifies

$$\begin{aligned}
&\|\mathcal{P}\mathbf{m} - \mathcal{P}\mathbf{m}_h\|_{L^2(\mathbb{R}^d)} + \|\nabla\phi^{**}(\mathbf{m}) - \nabla\phi^{**}(\mathbf{m}_h)\|_{L^2(\Omega)} \\
&\quad + \|\lambda\mathbf{m} - \lambda_h\mathbf{m}_h\|_{L^2(\Omega)} = \mathcal{O}(\varepsilon + h)
\end{aligned}$$

in the non-stabilized case. This suggests the choice $\varepsilon = h$ for the penalization parameter. Moreover, we see that the stabilization term on the right-hand side of (14) should satisfy at least $\sigma(\mathbf{n}_\mathcal{T}, \mathbf{n}_\mathcal{T}) = \mathcal{O}(h^2)$ for a smooth function $\mathbf{n} \in H^1(\Omega;\mathbb{R}^d)$.

The Galerkin orthogonality does not involve the term $(\mathbf{m} - \mathbf{m}) \cdot \mathbf{e}$, but only $\nabla\phi^{**}(\mathbf{m}) - \nabla\phi^{**}(\mathbf{m}_h)$. This is why the error estimates in [8,11,13] all lead to the term

$$\|\nabla\phi^{**}(\mathbf{m}) - \nabla\phi^{**}(\mathbf{m}_h)\|^2_{L^2(\Omega)} = \sum_{j=2}^{d} \|(\mathbf{m} - \mathbf{m}_h) \cdot \mathbf{z}_j\|^2_{L^2(\Omega)}$$

on the left-hand side and do not give control over $\|(\mathbf{m} - \mathbf{m}_h) \cdot \mathbf{e}\|_{L^2(\Omega)}$. (Here, we used the Definition of ϕ^{**} in the uniaxial case and Parseval's identity.)

While the cited result applies for $2D$ and $3D$, the following arguments are so far restricted to $2D$ only. One can prove that the error $\boldsymbol{\delta}_h = \mathbf{m} - \mathbf{m}_h$ in the $\widetilde{H}^1(\Omega)$-norm (i.e. the dual norm of $H^1(\Omega)$ with respect to the extended L^2-scalar product) satisfies

$$\begin{aligned}
&\|\mathbf{m} - \mathbf{m}_h\|_{\widetilde{H}^1(\Omega;\mathbb{R}^2)} \qquad (15) \\
&\quad \le C\left(\|\mathcal{P}\mathbf{m} - \mathcal{P}\mathbf{m}_h\|_{L^2(\mathbb{R}^2;\mathbb{R}^2)} + \|(\mathbf{m} - \mathbf{m}_h) \cdot \mathbf{z}_2\|_{L^2(\Omega)}\right).
\end{aligned}$$

Here, the terms on the right-hand side are controlled by (14). In particular, we obtain fully *weak* convergence of $\mathbf{m}_h$ towards $\mathbf{m}$ in $L^2(\Omega;\mathbb{R}^2)$, with the same order of convergence as before [10].

The second stabilization term on the left-hand side of (14) provides an additional convergence property that has not been used so far. Considering the stabilization (13), one can derive

$$\|\mathbf{m} - \mathbf{m}_h\|_{L^2(\Omega;\mathbb{R}^2)} = \mathcal{O}\big((h + \varepsilon)^{1/2}\big). \qquad (16)$$

Details will appear in [10]. Although this result is the only *full* L^2 convergence result for the proposed lowest order discretization, the numerical experiments give evidence that one can do better: even for the non-stabilized discretization we observe *full* L^2 convergence of optimal order $\mathcal{O}(h + \varepsilon)$.

2.4 A posteriori error control

So far, a reasonable a posteriori error analysis has only been established for the non-stabilized discretization. Using the notation from the previous section, we have

$$\begin{aligned}
&\|\mathcal{P}\mathbf{m} - \mathcal{P}\mathbf{m}_h\|^2_{L^2(\mathbb{R}^d)} + \|\nabla\phi^{**}(\mathbf{m}) - \nabla\phi^{**}(\mathbf{m}_h)\|^2_{L^2(\Omega)} \\
&\quad \le 2\big\{\|\varepsilon\,\lambda_h\mathbf{m}_h\|^2_{L^2(\Omega)} \\
&\qquad + \|\varepsilon|\lambda_h\mathbf{m}_h|\{(\mathbf{f} - \mathbf{f}_\mathcal{T}) - (\mathcal{P}\mathbf{m}_h - (\mathcal{P}\mathbf{m}_h)_\mathcal{T})\}\|_{L^1(\Omega)} \\
&\qquad + \langle(\mathbf{f} - \mathbf{f}_\mathcal{T}) - (\mathcal{P}\mathbf{m}_h - (\mathcal{P}\mathbf{m}_h)_\mathcal{T})\,;\,\mathbf{m} - \mathbf{m}_\mathcal{T}\rangle_{L^2}\big\}. \qquad (17)
\end{aligned}$$

The right-hand side of (17) is not fully computable as it contains $\mathbf{m} - \mathbf{m}_\mathcal{T}$. Since $|\mathbf{m}| \le 1$ there follows $\|\mathbf{m} - \mathbf{m}_\mathcal{T}\|_{L^\infty(\Omega)} \le 2$ and, hence, by Hölder's inequality

$$\langle(\mathbf{f}-\mathbf{f}_{\mathcal{T}})-(\mathcal{P}\mathbf{m}_h-(\mathcal{P}\mathbf{m}_h)_{\mathcal{T}})\,;\,\mathbf{m}-\mathbf{m}_{\mathcal{T}}\rangle_{L^2}$$
$$\leq 2\|(\mathbf{f}-\mathbf{f}_{\mathcal{T}})-(\mathcal{P}\mathbf{m}_h-(\mathcal{P}\mathbf{m}_h)_{\mathcal{T}})\|_{L^1(\Omega)}.$$

For $\mathbf{m}\in W^{1,\infty}(\Omega;\mathbb{R}^d)$, the Poincaré estimate leads to

$$\langle(\mathbf{f}-\mathbf{f}_{\mathcal{T}})-(\mathcal{P}\mathbf{m}_h-(\mathcal{P}\mathbf{m}_h)_{\mathcal{T}})\,;\,\mathbf{m}-\mathbf{m}_{\mathcal{T}}\rangle_{L^2}$$
$$\leq C\,\|h\{(\mathbf{f}-\mathbf{f}_{\mathcal{T}})-(\mathcal{P}\mathbf{m}_h-(\mathcal{P}\mathbf{m}_h)_{\mathcal{T}})\}\|_{L^1(\Omega)}$$

with a constant $C=C_P\,\|\mathbf{m}\|_{W^{1,\infty}(\Omega)}$.

The two resulting a posteriori error estimates are reliable (in the first case, as no assumption on the smoothness of the unknown solution $\mathbf{m}$ is included) but not efficient (as the estimate behaves too coarsely – cf. [8] and the numerical examples below); or, conversely, efficient (as the second estimate gives the higher convergence rates) but *not* reliable (or it appears doubtful to assume a higher order of smoothness such as $\mathbf{m}\in W^{1,\infty}(\Omega;\mathbb{R}^d)$). This phenomenon is called reliability-efficiency gap in [6]: What is reliable is not efficient and what is efficient is not reliable. The reason is again the lack of control over the term $\|\mathbf{m}-\mathbf{m}_h\|_{L^2(\Omega)}$.

We expect to overcome this drawback by an appropriate stabilization. This problem is topic of our current research.

3 Implementational aspects

This section presents the details on the numerical algorithms with emphasis on a globally convergent gradient method versus the classical Newton–Raphson solver. The affect of the small penalty parameter is studied in [9] and not discussed in detail here. It can be observed in numerical experiments with smooth solution that one has to choose $\varepsilon=\mathcal{O}(h)$.

3.1 Computation of discrete solution

To compute the unique solution $\mathbf{m}_h$ of $(RP_{\varepsilon,h})$ numerically, we consider a basis representation of $\mathbf{m}_h$. With respect to a triangulation $\mathcal{T}=\{T_1,\dots,T_N\}$ the discrete equations lead to the unknown coefficients $\mathbf{x}\in\mathbb{R}^{dN}$ of

$$\mathbf{m}_h=\sum_{j=1}^{N}\sum_{\alpha=1}^{d}x_{[j,\alpha]}\boldsymbol{\varphi}_{[j,\alpha]}\in L^\infty(\Omega;\mathbb{R}^d). \tag{18}$$

Here, one abbreviates $[j,1]:=j$ and $[j,2]:=j+N$ in $2D$ (and furthermore $[j,3]:=j+2N$ for $3D$) to fix the order of the coefficients and basis functions $\boldsymbol{\varphi}_{[j,\alpha]}:=\chi_{T_j}\mathbf{e}_\alpha$, where χ_{T_j} denotes the characteristic function of T_j and $\mathbf{e}_\alpha$ is the α-th standard unit vector. If we identify the coefficient vector $\mathbf{x}$ with $\mathbf{m}_h$, the discrete problem $(RP_{\varepsilon,h})$ is recast into a nonlinear system of equations

$$\mathbf{F}(\mathbf{x}):=\big(\langle\mathcal{P}\mathbf{m}_h+\nabla\phi^{**}(\mathbf{m}_h)+\lambda_h\mathbf{m}_h-\mathbf{f}\,;\,\boldsymbol{\varphi}_j\rangle_{L^2}$$
$$+\sigma(\mathbf{m}_h,\boldsymbol{\varphi}_j)\big)_{j=0}^{dN}=0\in\mathbb{R}^{dN} \tag{19}$$

and then solved by an iterative solver. In [8,9] we suggested the use of a (classical) Newton–Raphson scheme. But since F is not continuously differentiable (according to the definition of λ_h), convergence of the classical Newton–Raphson scheme is unclear.

3.2 A globally convergent gradient method

We recall an abstract result from [2] stated in the spirit of [24]: let X be a Hilbert space, and let $a_j(\cdot,\cdot)$ be a sequence of equivalent scalar products on X, i.e. there are constants $\alpha_j,M_j>0$ such that, for all $x,y\in X$,

$$\alpha_j\|x\|_X^2\leq a_j(x,x)\quad\text{and}\quad a_j(x,y)\leq M_j\|x\|_X\|y\|_X. \tag{20}$$

Let $E:X\to\mathbb{R}$ be a Gâteaux-differentiable energy functional on X such that there are constants $\beta,L>0$ with

$$\beta\|x-y\|_X^2+DE(x)(y-x)\leq E(y)-E(x), \tag{21}$$
$$\big(DE(x)-DE(y)\big)(x-y)\leq L\|x-y\|_X^2, \tag{22}$$

for all $x,y\in X$. Then, one can prove that the functional E has a unique minimizer $x^*\in X$. Given an arbitrary initial value $x_0\in X$ and a successively defined sequence (x_j),

$$a_j(x_j-x_{j+1},\cdot)=DE(x_j)\in X^*\quad\text{for } j\in\mathbb{N}, \tag{23}$$

one can prove convergence provided that $0<\beta+\alpha_j-L$, More precisely, there holds for $\delta_j:=E(x_j)-E(x^*)\geq 0$

$$\delta_{j+1}\leq(1-4\beta(\beta+\alpha_j-L)M_j^{-2})\delta_j, \tag{24}$$
$$\|x^*-x_j\|_X^2\leq M_j^2\beta^{-2}(\beta+\alpha_j-L)^{-1}(\delta_j-\delta_{j+1}). \tag{25}$$

In particular, the energy sequence $E(x_j)$ is monotonously decreasing. Provided there are constants $\alpha_0,M_0>0$ such that $\alpha_0\leq\alpha_j\leq M_j\leq M_0$ for all $j\in\mathbb{N}$, one obtains convergence $\lim_{j\to\infty}x_j=x^*$ in X. We apply this result to compute the unique solution of $(\mathrm{RP}_{\varepsilon,h})$, where we consider the stabilization (12) and $\beta=h_{\min}^2/2$ with the minimal mesh-size $h_{\min}:=\min h$.

Corollary 1 *Let $\mathbf{m}_h^{(0)}\in\mathcal{L}^0(\mathcal{T})^d$ be an arbitrary initial value, and define a sequence of magnetizations $\mathbf{m}_h^{(j)}$ in $\mathcal{L}^0(\mathcal{T})^d$ inductively for $j\in\mathbb{N}$ by*

$$\mathbf{m}_h^{(j+1)}:=\mathbf{m}_h^{(j)}-\frac{1}{\alpha}\Big\{\Big(\mathcal{P}\mathbf{m}_h^{(j)}\Big)_{\mathcal{T}}+\nabla\phi^{**}\Big(\mathbf{m}_h^{(j)}\Big)$$
$$+\lambda_h^{(j)}\mathbf{m}_h^{(j)}-\mathbf{f}_{\mathcal{T}}+h^2\mathbf{m}_h^{(j)}\Big\} \tag{26}$$

with any $\alpha > 2 + 3/\varepsilon_{\min}$ *and* $\varepsilon_{\min} = \min \varepsilon$. *Then,* $(\mathbf{m}_h^{(j)})$ *converges to the unique solution* $\mathbf{m}_h$ *of* $(RP_{\varepsilon,h})$ *stabilized by* (12), *i.e.* $\sigma(\mathbf{m}_h, \mathbf{n}_h) = \langle h\mathbf{m}_h \,;\, h\mathbf{n}_h\rangle_{L^2}$.

Proof With respect to the abstract notation, we have $X = \mathcal{L}^0(\mathcal{T})^d$ equipped with the usual L^2 norm and

$$\begin{aligned} E(\mathbf{m}_h) = {} & \frac{1}{2}\int_{\mathbb{R}^d} |\mathcal{P}\mathbf{m}_h|^2 \,\mathrm{d}x + \frac{1}{2}\int_{\Omega} |\phi^{**}(\mathbf{m}_h)|^2 \,\mathrm{d}x \\ & - \int_{\Omega} \mathbf{f}\cdot\mathbf{m}_h \,\mathrm{d}x + \frac{1}{2\varepsilon}\int_{\Omega} (|\mathbf{m}_h| - 1)_+^2 \,\mathrm{d}x \\ & + \frac{1}{2}\sigma(\mathbf{m}_h, \mathbf{m}_h). \end{aligned} \tag{27}$$

Note that E is convex and (Fréchet-) differentiable with derivative

$$\begin{aligned} DE(\mathbf{m}_h)(\mathbf{n}_h) = {} & \langle \mathcal{P}\mathbf{m}_h + \nabla\phi^{**}(\mathbf{m}_h) - \mathbf{f} + \lambda_h\mathbf{m}_h \,;\, \mathbf{n}_h\rangle_{L^2} \\ & + \sigma(\mathbf{m}_h, \mathbf{n}_h) \end{aligned}$$

for all $\mathbf{n}_h \in \mathcal{L}^0(\mathcal{T})^d$. Therefore, $(\mathrm{RP}_{\varepsilon,h})$ is equivalent to $DE(\mathbf{m}_h) = 0$ in $\mathcal{L}^0(\mathcal{T};\mathbb{R}^d)^*$, i.e. $x^* = \mathbf{m}_h$. To verify (21) for $\beta = h_{\min}^2/2$, we check the estimate for each energy contribution separately. Obviously, for scalars $a, b \in \mathbb{R}$ there holds $2a(b-a) \le (a-b)^2 + 2a(b-a) = b^2 - a^2$. If we apply this elementary calculation to the occurring bilinear forms, we obtain

$$\begin{aligned} & \langle \mathcal{P}\mathbf{m}_h \,;\, \mathbf{n}_h - \mathbf{m}_h\rangle_{L^2} \\ & \quad \le \frac{1}{2}\|\mathcal{P}\mathbf{n}_h\|^2_{L^2(\mathbb{R}^d)} - \frac{1}{2}\|\mathcal{P}\mathbf{m}_h\|^2_{L^2(\mathbb{R}^d)}, \\ & \quad \langle \nabla\phi^{**}(\mathbf{m}_h) \,;\, \mathbf{n}_h - \mathbf{m}_h\rangle_{L^2} \\ & \quad \le \frac{1}{2}\|\phi^{**}(\mathbf{n}_h)\|^2_{L^2(\Omega)} - \frac{1}{2}\|\phi^{**}(\mathbf{m}_h)\|^2_{L^2(\Omega)}, \\ & \quad \beta\|\mathbf{m}_h - \mathbf{n}_h\|^2_{L^2} + \sigma(\mathbf{m}_h, \mathbf{n}_h - \mathbf{m}_h) \\ & \quad \le \frac{1}{2}\sigma(\mathbf{m}_h - \mathbf{n}_h, \mathbf{m}_h - \mathbf{n}_h) + \sigma(\mathbf{m}_h, \mathbf{n}_h - \mathbf{m}_h) \\ & \quad = \frac{1}{2}\sigma(\mathbf{n}_h, \mathbf{n}_h) - \frac{1}{2}\sigma(\mathbf{m}_h, \mathbf{m}_h). \end{aligned}$$

Since the third term in (27) is linear, it remains to estimate the penalization contribution. This follows from the pointwise estimate

$$2\,\frac{(|x|-1)_+}{|x|}\, x\cdot(y-x) \le (|y|-1)_+^2 - (|x|-1)_+^2$$

for all $x, y \in \mathbb{R}^d$, which is easily checked by direct calculations. To verify the Lipschitz condition (22) with $L = 2 + 3/\varepsilon_{\min} + \beta$ we proceed analogously. The only point of interest is the Lipschitz continuity of the mapping $\Lambda : \mathcal{L}^0(\mathcal{T})^d \to \mathcal{L}^0(\mathcal{T})^d, \mathbf{m}_h \mapsto \lambda_h\mathbf{m}_h$ with respect to the L^2 norm and Lipschitz constant $3/\varepsilon_{\min}$. But this follows from the pointwise estimate

$$\left|\frac{(|x|-1)_+}{|x|}\, x - \frac{(|y|-1)_+}{|y|}\, y\right| \le 3\,|x-y|$$

for all $x, y \in \mathbb{R}^d$, which is again obtained by a direct calculation. Note that α satisfies $0 < \beta + \alpha - L$. Now, consider the scalar product $a_j(\mathbf{m}_h, \mathbf{n}_h) = \alpha\langle \mathbf{m}_h \,;\, \mathbf{n}_h\rangle_{L^2}$ and $\alpha_j = \alpha = M_j$ so that (20) is satisfied. According to (23) the inductive definition of $\mathbf{m}_h^{(j)}$,

$$\begin{aligned} & \alpha\langle \mathbf{m}_h^{(j)} - \mathbf{m}_h^{(j+1)} \,;\, \mathbf{n}_h\rangle_{L^2} \\ & \quad = \langle \mathcal{P}\mathbf{m}_h^{(j)} + \nabla\phi^{**}(\mathbf{m}_h^{(j)}) + \lambda_h^{(j)}\mathbf{m}_h^{(j)} - \mathbf{f} \,;\, \mathbf{n}_h\rangle_{L^2} \\ & \qquad + \langle h\mathbf{m}_h \,;\, h\mathbf{n}_h\rangle_{L^2}, \end{aligned}$$

for all $\mathbf{n}_h \in \mathcal{L}^0(\mathcal{T})^d$, defines a convergent sequence with limit $\mathbf{m}_h$. Recalling that the mapping $\mathbf{m} \mapsto \mathbf{m}_{\mathcal{T}}$ is the L^2 projection onto $\mathcal{L}^0(\mathcal{T})^d$, we can exchange $\mathcal{P}\mathbf{m}_h^{(j)}$ and $\mathbf{f}$ by $(\mathcal{P}\mathbf{m}_h^{(j)})_{\mathcal{T}}$ and $\mathbf{f}_{\mathcal{T}}$, respectively. Testing the last equation with the basis functions yields (26). □

The computation of the L^2 scalar product

$$A_{[j,\alpha][k,\beta]} := \langle \mathcal{P}\boldsymbol{\varphi}_{[j,\alpha]} \,;\, \boldsymbol{\varphi}_{[k,\beta]}\rangle_{L^2}$$

required for the evaluation of $\mathbf{F}(x)$ in (19) is possible with a closed form formula,

$$\begin{aligned} A_{[j,\alpha][k,\beta]} := {} & \int_{\mathbb{R}^d} (\mathcal{P}\boldsymbol{\varphi}_{[k,\beta]})\cdot\boldsymbol{\varphi}_{[j,\alpha]} \,\mathrm{d}x \\ = {} & -\int_{\partial T_j}\int_{\partial T_k} G(x-y)\nu_\alpha^{(j)}(x)\nu_\beta^{(k)}(y) \,\mathrm{d}s_y \,\mathrm{d}s_x \end{aligned} \tag{28}$$

(with outer normal vectors $\boldsymbol{\nu}^{(j)}$ and $\boldsymbol{\nu}^{(k)}$ on the boundaries ∂T_j and ∂T_k, respectively), cf. [27]. The computation of the double boundary integral can be performed exactly [7,15,23]. In particular, $(\mathcal{P}\mathbf{m}_h)_{\mathcal{T}}$ occurring in Corollary 1 can be computed analytically.

Remark 3 It is possible to apply $\mathcal{H}$-matrix techniques [3] to treat the matrix $\mathbf{A}$ efficiently [8,9,16,26,27]. This allows to reduce the numerical cost for both storage and matrix–vector multiplication down to (almost) linear.

The factor $\lambda := 1/\alpha$ in (26) can be interpreted as a damping factor for the step size, where the search direction is given by the gradient of the energy functional. A priori α has to be relatively large and this is also confirmed by the numerical experiments. Even if we combine the introduced gradient method with a line search method, the number of iterations remains to be a couple of thousands even on a uniform mesh with less than 500 elements. On the other hand, as has been studied in [9], the classical Newton–Raphson scheme behaves perfectly and converges in less than 20 iterations. This makes the above algorithm impractical for real-life experiments. For the numerical experiments in [8,9], we used a classical Newton–Raphson scheme with initial value obtained by prolongation of coarse grid solutions.

The corresponding algorithm can be stated quite efficiently. It is shown in [9] that – up to one matrix-vector multiplication with $\mathbf{A}$ – the computation of residual $\mathbf{F}(x)$ and Jacobian $D\mathbf{F}(x)$ can be performed in $\mathcal{O}(N)$ operations, where N denotes the number of elements.

3.3 Adaptive mesh-refinement

The nested Newton–Raphson solver is part of a multilevel scheme driven by an adaptive algorithm. The a posteriori estimate (17) gives rise to the error estimators

$$\mu := \left(\sum_{T\in\mathcal{T}} \mu_T^2\right)^{\frac{1}{2}} \quad \text{and} \quad \eta := \left(\sum_{T\in\mathcal{T}} \eta_T^2\right)^{\frac{1}{2}} \tag{29}$$

with the refinement indicators μ_T, η_T, for $T \in \mathcal{T}$, defined by

$$\begin{aligned} \ell_T &:= (\varepsilon\lambda_h|\mathbf{m}_h|)|_T = \max\{0, |\mathbf{m}_h|_T| - 1\}, \\ \mu_T^2 &:= (1+\ell_T)\|(\mathbf{f}-\mathbf{f}_{\mathcal{T}}) - (\mathcal{P}\mathbf{m}_h - (\mathcal{P}\mathbf{m}_h)_{\mathcal{T}})\|_{L^1(T)} \\ &\quad + |T|\ell_T^2, \\ \eta_T^2 &:= (h_T+\ell_T)\|(\mathbf{f}-\mathbf{f}_{\mathcal{T}}) - (\mathcal{P}\mathbf{m}_h - (\mathcal{P}\mathbf{m}_h)_{\mathcal{T}})\|_{L^1(T)} \\ &\quad + |T|\ell_T^2. \end{aligned} \tag{30}$$

The estimator μ is reliable, i.e. an upper bound for the error $\|\mathcal{P}\mathbf{m} - \mathcal{P}\mathbf{m}_h\|_{L^2(\mathbb{R}^d)} + \|\nabla\phi^{**}(\mathbf{m}) - \nabla\phi^{**}(\mathbf{m}_h)\|_{L^2(\Omega)}$ up to a multiplicative constant, but cannot be efficient, i.e also a lower bound; the estimator η is reliable solely for $\mathbf{m} \in W^{1,\infty}(\Omega;\mathbb{R}^d)$, but expected to be efficient.

Algorithm 2 (Adaptive Mesh-Refinement) Input: *Initial triangulation $\mathcal{T}^{(0)}$, $\alpha > 0$, and $0 \le \theta \le 1$. Set $n = 0$ and $\mathbf{m}_h := 0$.*
Compute for $n = 1, 2, \ldots$ until termination
(i) *On $T_j \in \mathcal{T}^{(n)} = \{T_1, \ldots, T_N\}$ set $\varepsilon|_{T_j} := \varepsilon_j = h_{T_j}^{\alpha} > 0$, $j = 1, \ldots, N$*
(ii) *Call Newton-Raphson scheme with start vector associated to $\mathbf{m}_h^{(n-1)}$ and output $\mathbf{m}_h^{(n)}$.*
(iii) *Compute μ and η from (29) and indicators $\eta_j := \eta_{T_j}$ and $\mu_j := \mu_{T_j}$ from (30) with $\mathbf{m}_h$ substituted by $\mathbf{m}_h^{(n)}$.*
(iv) *Mark an element $T_j \in \mathcal{T}^{(n)}$ provided $\eta_j \ge \theta \max_{1\le k\le N} \eta_k$ (or μ_j respectively).*
(v) *Refine the marked elements, update n and go to (i).*
Output: *Sequence of $\mathcal{T}^{(n)}$, $\eta^{(n)}$, $\mu^{(n)}$, $\mathbf{m}_h^{(n)}$ for $n = 1, 2, \ldots$.*

The choice $\theta = 0$ in Algorithm 2 leads to uniform mesh-refinement, whereas $\theta = 1/2$ leads to adapted meshes. The following example underlines that adaptive mesh-refinement can improve the order of convergence up to the optimal order $\mathcal{O}(h)$ in terms of the mesh-size h.

To study the performance of the mesh-adaptation, we consider an example from [8] with known *non-smooth* exact solution $(\lambda, \mathbf{m})$ on the unit square $\Omega = (0,1)^2$,

Table 1 Experimental results in Sect. 3.3 for uniform (top), μ-adaptive, and η-adaptive (bottom) mesh-refinement with penalization parameter $\varepsilon = h$.

k	$N^{(k)}$	$n^{(k)}$	$E^{(k)}$	$\kappa^{(k)}(E)$	$\eta^{(k)}$	$\kappa^{(k)}(\eta)$	$\eta^{(k)}/E^{(k)}$	$\mu^{(k)}$	$\kappa^{(k)}(\mu)$	$E^{(k)}/\mu^{(k)}$
					Uniform mesh-refinement					
0	4	8	0.463		0.660		1.425	0.749		0.618
1	16	7	0.275	0.376	0.361	0.435	1.312	0.542	0.233	0.507
2	64	11	0.158	0.401	0.186	0.480	1.176	0.393	0.233	0.402
3	256	10	0.094	0.370	0.089	0.527	0.947	0.274	0.259	0.344
4	1024	12	0.056	0.381	0.044	0.510	0.792	0.196	0.241	0.284
5	4096	12	0.032	0.394	0.023	0.481	0.701	0.146	0.215	0.222
					μ-Adaptive mesh-refinement					
0	4	8	0.463		0.660		1.425	0.749		0.618
3	106	9	0.122	0.537	0.144	0.553	1.185	0.327	0.285	0.373
5	331	11	0.071	0.549	0.072	0.737	1.008	0.235	0.317	0.304
7	892	11	0.045	0.373	0.048	0.469	1.063	0.184	0.256	0.246
9	4948	15	0.019	0.447	0.018	0.487	0.936	0.120	0.209	0.162
					η-Adaptive mesh-refinement					
0	4	8	0.463		0.660		1.425	0.749		0.618
3	100	10	0.128	0.471	0.144	0.575	1.127	0.333	0.362	0.384
5	373	9	0.068	0.765	0.068	0.578	0.995	0.229	0.333	0.299
7	1183	10	0.043	0.803	0.036	0.846	0.838	0.171	0.358	0.249
9	4516	14	0.023	0.348	0.018	0.522	0.801	0.128	0.160	0.181

The table shows the outcome of Algorithm 2 with respect to the k-th mesh $\mathcal{T}_k$: the number of elements N, the number of steps n in the Newton–Raphson scheme to compute $\mathbf{m}_h$, the error $E = \|\mathbf{m} - \mathbf{m}_h\|_{L^2(\Omega)}$, and the error estimators η and μ. Furthermore, the table shows the experimental convergence rates $\kappa^{(k)}$ of the error E and the estimators η, μ, and the quotients η/E and E/μ. In contrast to the case of uniform mesh-refinement, for the adaptive schemes the efficiency quotient η/E seems to converge to a value close to 1. The reliability quotient E/μ tends to zero as for uniform mesh-refinement

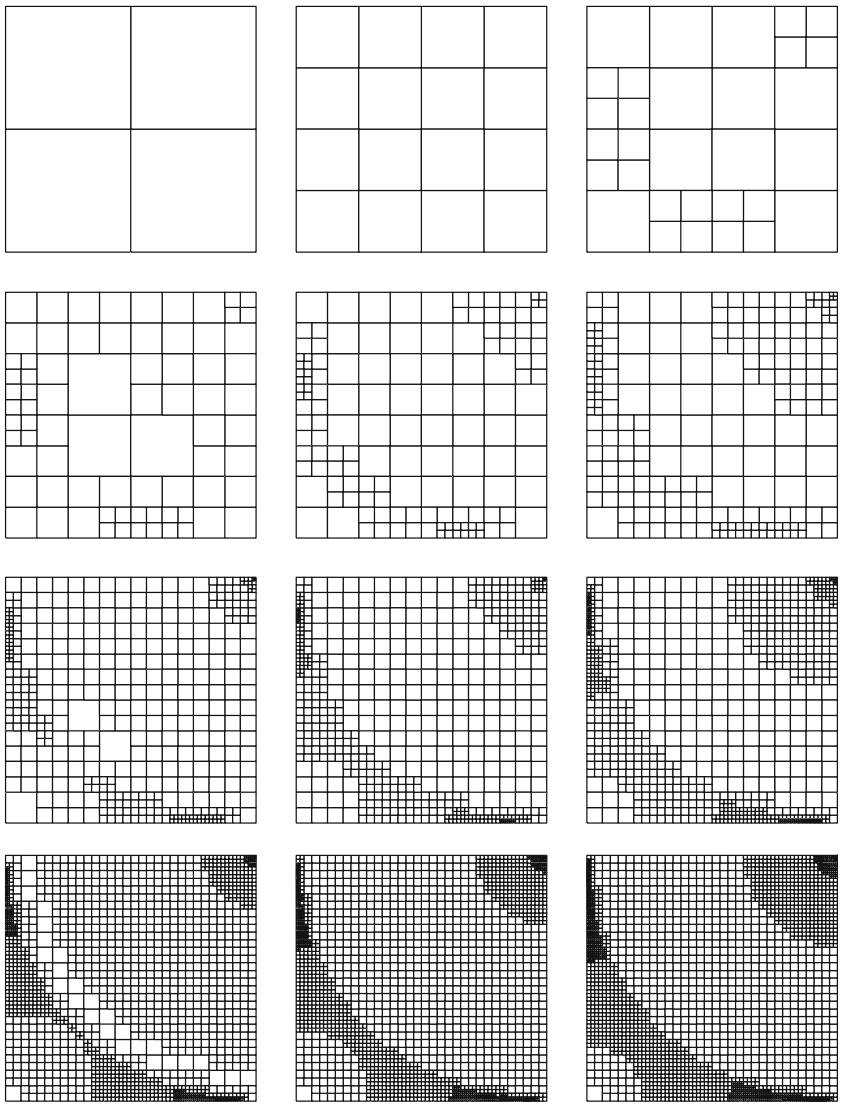

Fig. 1 Sequences of adaptively generated meshes $\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_{12}$ by Algorithm 2 with refinement indicator ϱ (with $N = 4, 16, 31, 76, 142, 208, 415, 586, 911, 1, 717, 2, 245, 2, 779$). The meshes are highly adapted towards the right upper corner $(1, 1)$ and along the interface $|(1, 1) - x| = 1$

$$(\mathbf{m}(x), \lambda(x)) := \begin{cases} (\mathbf{y}(x), 0) & \text{for } x \in \omega, \\ \frac{x_1}{1-y_1(x)} \, \frac{x_2}{1-y_2(x)} \, \mathbf{y}(x), 1) & \text{for } x \in \Omega \backslash \omega \end{cases} \tag{31}$$

with a singular gradient at the three vertices $(0, 1)$, $(1, 0)$, $(1, 1)$ on the boundary of the magnetic body $\Omega = (0, 1)^2$ leading to $\mathbf{m} \notin H^1(\Omega; \mathbb{R}^2)$. Here,

$$\mathbf{y}(x) := \frac{(1, 1) - x}{|(1, 1) - x|} \quad \text{and} \quad \omega := \big\{ x \in \Omega \,\big|\, |(1, 1) - x| < 1 \big\}.$$

The unit square Ω is filled with a uni-axial magnetic material with easy axis $\mathbf{e} = (-1, 1)/\sqrt{2}$, i.e. $\mathbf{z} = (1, 1)/\sqrt{2}$ in Eq. (4). The errors $\|\mathbf{m} - \mathbf{m}_h\|_{L^2(\Omega)}$, the error estimators η and μ, and their experimental convergence rates κ in terms of the number N of degrees of freedom are compared in Table 1 for uniform, μ- and η-adaptive mesh-refinements. The quotient E/μ is bounded from above and decreases. This is in agreement with the reliability and the expected overestimation (owing to the reliability-efficiency- gap). The quotients η/E are motivated by an expected efficiency of η and they are, in fact, bounded from above according to this. The convergence rates for the error E are strongly improved by adaptive mesh-refining in contrast to the convergence rates for the error estimators η and μ. Finally, the reliability-efficiency-gap is reflected explicitly by the quite different convergence rates κ for η and μ.

Figure 1 shows meshes that are obtained from our adaptive algorithm with the refinement indicator $\varrho_{T_j} := \|\mathbf{m} - \mathbf{m}_{\mathcal{T}}\|_{L^2(T_j)}$. Although there is no proof, we expect that the ϱ-adaptively generated meshes are somehow optimal since they try to equidistribute the best-approximation error over Ω. Figure 2 displays meshes obtained by η-adaptive mesh-refinement. Note that the generated meshes are quite similar to the *optimal* meshes. This is also reflected in the experimental convergence order of

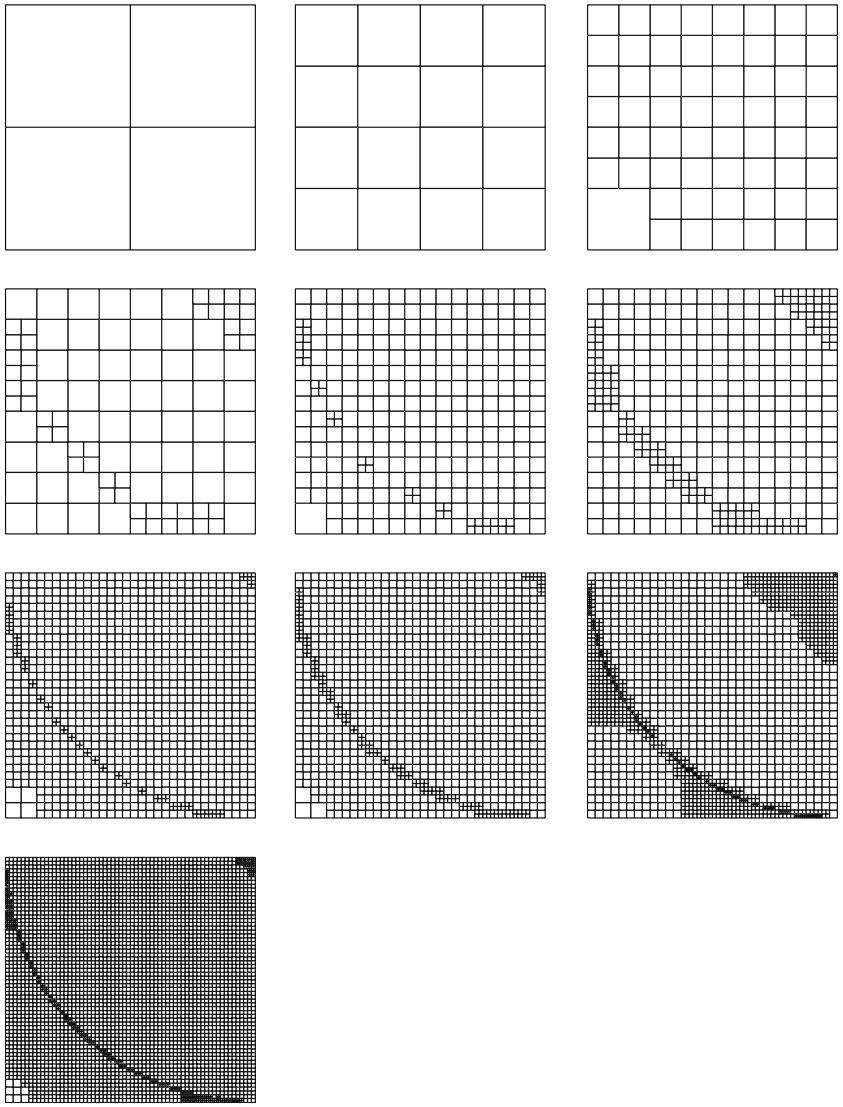

Fig. 2 Sequences of adaptively generated meshes $\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_9$ (with $N = 4, 16, 61, 100, 286, 373, 1, 114, 1, 183, 1, 912, 4, 516$) obtained by Algorithm 2 with refinement indicator η and penalization parameter $\varepsilon = h$. Notice that we obtain meshes which are close to the ϱ-adaptively generated meshes of Fig. 1

the error in Table 1. Altogether, we obtain the optimal convergence order $\mathcal{O}(N^{1/2})$ corresponding to $\mathcal{O}(h)$ for $2D$.

4 Conclusions

The final section presents conclusions, comments, and remarks on future developments.

4.1 Resume

In the large and soft body limit of micromagnetics, the effective magnetization vector, i.e. the space average of the micromagnetic magnetization vectors, can be calculated directly from an effective model (RP). Therein, the exterior field problem can be recast via some Helmholtz projection operator $\mathcal{P}$ that allows a nonlocal problem for L^∞ functions on the magnetic domain Ω. The convexified pointwise side-restriction $|\mathbf{m}| \leq 1$ can be involved in a penalization strategy. The associated discrete problem ($\mathrm{RP}_{\varepsilon,h}$) acts on piecewise constant trial and test functions. The work [8] presents a throughout a priori and a posteriori error analysis of the discretization errors which is sharpened in [10].

Iterative schemes for the solution of the linear subproblems occurring by the solution of ($\mathrm{RP}_{\varepsilon,h}$) are supported by an $\mathcal{H}$-matrix approach [3,16,26].

4.2 Effective modelling of effective magnetization

The model ($\mathrm{RP}_{\varepsilon,h}$) allows the efficient simulation of the effective magnetization vectors. This yields a macroscopic approximation of a multi-scale problem with

a complicated microscopic structure which would be impossible to compute by a resolution of the fines scale phenomena. We refer to [5] for a one-dimensional trivial example of non-convex minimization problem that illustrates that cluster of local minimizers in the high-dimensional global non-convex minimization problem yield an extremely difficult discrete problem one should not assume to be able to solve accurately.

4.3 Stabilization and penalization

The penalty parameter $\varepsilon = h^\alpha$ for $\alpha > 1$ small, such as $\varepsilon = h$ seems to be a good compromise between accuracy (α large) and the condition of the discrete system of equations (α small). In contrast to other situations in convexified problems [1], a further stabilization is not necessary but might yield further convergence properties of the scheme.

4.4 Future developments

The a posteriori error estimates in this paper show the reliability-efficiency gap for the error estimators μ (proven to be reliable) and η (expected to be efficient). This dramatic lack of error control requires to be overcome in the future for reliable and accurate numerical simulations. In the numerical experiments one obtains *full* L^2 convergence which is not yet understood and has to be the topic of further research. Finally, establishing rigorous convergence properties of iterative schemes for the solution of $(\mathrm{RP}_{\varepsilon,h})$ might be an interesting task.

Acknowledgements The support of the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin and the Austrian Science Fund FWF under grant P15274 is thankfully acknowledged.

References

1. Bartels S, Carstensen C, Plechac P, Prohl A (2004) Convergence for stabilisation of degenerately convex minimisation problems. J Interface Free Bound Probl 6(2):253–269
2. Bartels S, Carstensen C, Hackl K, Hoppe U (2003) Effective relaxation for microstructure simulations: algorithms and applications. Comput Methods Appl Mech Eng 193:5143–5175
3. Börm S, Grasedyck L, Hackbusch W (2003) Introduction to hierarchical matrices with applications. Eng Anal Bound Elem 27:405–422
4. Brown W (1963) Micromagnetics. Wiley, New York
5. Carstensen C (2001) Numerical analysis of microstructure. In: Blowey JF, Coleman JP, Craig AW (eds) Chapter II of Theory and Numerics of Differential Equations, Durham 2000, pp 59–126. Springer, Berlin Heidelberg, New York
6. Carstensen C, Jochimsen K (2003) Adaptive finite element error control for non-convex minimization problems: numerical two-well model example allowing microstructures. Computing 71:175–204
7. Carstensen C, Praetorius D (2004) A posteriori error control in adaptive qualocation boundary element analysis for a logarithmic-kernel integral equation of the first kind. SIAM J Sci Comp 25:259–283
8. Carstensen C, Praetorius D (2005) Numerical analysis for a macroscopic model in micromagnetics. SIAM J Numer Anal 42:2633–2651
9. Carstensen C, Praetorius D (2005) Effective simulation of a macroscopic model for stationary micromagnetics. Comput Methods Appl Mech Eng 194:531–548
10. Carstensen C, Praetorius D (2006) Stabilization yields strong convergence of macroscopic magnetization vectors for micromagnetics without exchange energy. IMA J Numer Anal (accepted)
11. Carstensen C, Prohl A (2001) Numerical analysis of relaxed micromagnetics by penalised finite elements. Numer Math 90:65–99
12. DeSimone A (1993) Energy minimizers for large ferromagnetic bodies. Arch Ration Mech Anal 125:99–143
13. Funken S, Prohl A (2005) Stabilization methods in relaxed micromagnetism. ESAIM Math Model Numer Anal 39:995–1017
14. Gilbarg D, Trudinger N (1977) Elliptic partial differential equations of second order. Springer, Berlin Heidelberg New York, 516 pp
15. Hackbusch W (2003) Direct integration of the Newton potential over cubes including a program description. Computing 68:193–216
16. Hackbusch W, Melenk J (2003) $\mathcal{H}$-matrix treatment of the operator $\nabla\Delta^{-1}\mathrm{div}$. (Unpublished Preprint)
17. Hubert A, Schäfer R (1998) Magnetic domains. Springer, Berlin Heidelberg New York xxiii+696
18. James R, Kinderlehrer D (1990) Frustration in ferromagnetic materials. Continuum Mech Thermondyn 2:215–239
19. Landau L, Lifshitz E (1965) On the theory of the dispersion of magnetic permeability in ferromagnetic bodies. Collected papers of L.D. Landau. Pergamon, New York pp 101–114
20. Lifshitz E (1944) On the magnetic structure of iron. J Phys USSR 8:337–346
21. Luskin M, Ma L (1992) Analysis of the finite element approximation of microstructure in micromagnetics. SIAM J Numer Anal 29:320–331
22. Ma L (1991) Analysis and computation for a variational problem in micromagnetics. PhD thesis, University of Minnesota
23. Maischak M (1999/2000) The analytical computation of the Galerkin elements for the Laplace, Lamé and Helmholtz Equation in BEM. Preprint 1999: *2D BEM*, Preprint 2000: *3D BEM*, Institut für Angewandte Mathematik, Universität Hannover
24. Nečas J (1986) Introduction to the theory on nonlinear elliptic equations. Wiley, Chichester, 1986
25. Pedregal P (1997) Parametrized measures and variational principles. Birkhäuser, Basel, 212 pp
26. Popović N, Praetorius D (2004) Applications of $\mathcal{H}$-matrix techniques in micromagnetics. Computing 74:177–204
27. Praetorius D (2004) Analysis of the Operator $\Delta^{-1}\mathrm{div}$ arising in magnetic models. Z Anal Anwendungen 23:589–605
28. Prohl A (2001) Computational micromagnetism. Teubner, Stuttgart, 304 pp
29. Tartar L (1995) Beyond young measures. Meccanica 30:505–526