Heterogeneous multiscale methods for the Landau-Lifshitz equation

Lena Leitenmaier[1*] and Olof Runborg[1]

[1]Department of Mathematics, KTH, Royal Institute of Technology, Lindstedtsvägen 25, Stockholm, 10044, Sweden.

*Corresponding author(s). E-mail(s): lenalei@kth.se;
Contributing authors: olofr@kth.se;

Abstract

In this paper, we present a finite difference Heterogeneous Multiscale Method for the Landau-Lifshitz equation with a highly oscillatory diffusion coefficient. The approach combines a higher order discretization and artificial damping in the so-called micro problem to obtain an efficient implementation. The influence of different parameters on the resulting approximation error is discussed. Further important factors that are taken into account are the choice of time integrator and the initial data for the micro problem which has to be set appropriately to get a consistent scheme. Numerical examples in one and two space dimensions and for both periodic as well as more general coefficients are given to demonstrate the functionality of the approach.

Keywords: Heterogeneous Multiscale Methods, Micromagnetics, Finite differences, Landau-Lifshitz

1 Introduction

The simulation of ferromagnetic composites can play an important role in the development of magnetic materials. A typical approach to describing magnetization dynamics of ferromagnetic materials is using the micromagnetic version of the Landau-Lifshitz equation, which states

$$\partial_t \mathbf{M}^\varepsilon = -\mathbf{M}^\varepsilon \times \mathbf{H}(\mathbf{M}^\varepsilon) - \alpha \mathbf{M}^\varepsilon \times [\mathbf{M}^\varepsilon \times \mathbf{H}(\mathbf{M}^\varepsilon)]\,, \tag{1}$$

where $\mathbf{M}^\varepsilon$ is the magnetization vector, $\mathbf{H}(\mathbf{M}^\varepsilon)$ the effective field acting on the magnetization and the material constant α describes the strength of damping. While the effective field contains several important contributions, we here consider a simplified model, only taking into account exchange interaction, and introduce a coefficient a^ε describing the material variations in the composite. The parameter $\varepsilon \ll 1$ represents the scale of these variations. We then have

$$\mathbf{H}(\mathbf{M}^\varepsilon) := \boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla} \mathbf{M}^\varepsilon),$$

a model that was first described in [1]. Similar models have also recently been used by for example [2] as well as [3] and [4].

For small values of ε, it becomes computationally very expensive and at some point infeasible to provide proper numerical resolution for a simulation of eq. (1). Hence we aim to apply numerical homogenization based on the approach of Heterogeneous Multiscale Methods (HMM) to the problem. In this framework, one combines a coarse scale macro problem with a micro problem resolving the relevant fast scales on a small domain in order to obtain an approximation to the effective solution corresponding to the problem.

For a simplified Landau-Lifshitz problem with a highly oscillatory external field and no spatial interaction, a possible HMM setup was introduced in [5] and extended to a non-zero temperature scenario in [6].

For the problem we consider here, eq. (1), the homogenization error has been analyzed in [7]. There it is also shown that $\mathbf{M}^\varepsilon$ exhibits fast oscillations in both space and time, where the spatial variations are of order $\mathcal{O}(\varepsilon)$ while the temporal ones are of order $\mathcal{O}(\varepsilon^2)$ and get damped away exponentially with time, depending on the given value of α. In [8] several ways to set up HMM were discussed and the errors introduced in the numerical homogenization process, the so-called upscaling errors, where analyzed. In this paper, we focus on numerical aspects related to the implementation of HMM for eq. (1). In Section 2, we first give an overview of the method and include relevant known results from [7] and [8]. We then discuss some aspects of time integration of the Landau-Lifshitz equation in Section 3 and suggest suitable methods for the time stepping in the macro and micro problem, respectively. Section 4 focuses on the HMM micro problem. We study how to choose initial data for the micro problem that is appropriately coupled to the current macro scale solution in Section 4.1, before using numerical example problems to investigate several factors that influence the errors introduced in the HMM averaging process in Section 4.3. In Section 5 we present numerical examples to show that the HMM approach can also be applied to locally-periodic and quasi-periodic problems.

2 Heterogeneous Multiscale Methods

In this section, we introduce the concept of Heterogeneous Multiscale Methods, discuss how we choose to set up a HMM model for the Landau-Lifshitz problem eq. (1) and give relevant error estimates that were introduced in [7] and [8].

2.1 Problem description

The specific multiscale problem we consider in this article is to find $\mathbf{M}^\varepsilon$ that satisfies the nonlinear initial value problem

$$\partial_t \mathbf{M}^\varepsilon = -\mathbf{M}^\varepsilon \times \boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla} \mathbf{M}^\varepsilon) - \alpha \mathbf{M}^\varepsilon \times [\mathbf{M}^\varepsilon \times \boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla} \mathbf{M}^\varepsilon)], \tag{2a}$$

$$\mathbf{M}^\varepsilon(x, 0) = \mathbf{M}_{\text{init}}(x), \tag{2b}$$

with periodic boundary conditions, on a fixed time interval $[0, T]$ and a spatial domain $\Omega = [0, L]^d$ for some $L \in \mathbb{R}$ and dimension $d = 1, 2, 3$. Here a^ε is a material coefficient which oscillates with a frequency determined by ε. We furthermore assume the following.

(A1) The material coefficient function a^ε is in $C^\infty(\Omega)$ and bounded by constants $a_{\min}, a_{\max} > 0$; it holds that $a_{\min} \leq a^\varepsilon(x) \leq a_{\max}$ for all $x \in \Omega$.

(A2) The damping coefficient α and the oscillation period ε are small, $0 < \alpha \leq 1$ and $0 < \varepsilon < 1$.

(A3) The initial data $\mathbf{M}_{\text{init}}(x)$ is normalized such that $|\mathbf{M}_{\text{init}}(x)| = 1$ for all $x \in \Omega$, which implies that $|\mathbf{M}^\varepsilon(x, t)| = 1$ for all $x \in \Omega$ and $t \in [0, T]$.

When the material coefficient is periodic, $a^\varepsilon = a(x/\varepsilon)$ where $a(y)$ is a 1-periodic function and $\varepsilon = L/\ell$ for some $\ell \in \mathbb{N}$, one can analytically derive a homogenized problem corresponding to eq. (2), as shown in [7]. The solution $\mathbf{M}_0$ to this homogenized problem satisfies

$$\partial_t \mathbf{M}_0 = -\mathbf{M}_0 \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{M}_0 \mathbf{A}^H) - \alpha \mathbf{M}_0 \times \mathbf{M}_0 \times \left[\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{M}_0 \mathbf{A}^H)\right], \tag{3a}$$

$$\mathbf{M}_0(x, 0) = \mathbf{M}_{\text{init}}, \tag{3b}$$

where $\mathbf{A}^H$ is the same homogenized coefficient matrix as for standard elliptic homogenization problems,

$$\mathbf{A}^H := \int_{[0,1]^d} a(y) \left(\mathbf{I} + (\boldsymbol{\nabla}_y \boldsymbol{\chi})^T\right) dy. \tag{4}$$

Here $\boldsymbol{\chi}(y) \in \mathbb{R}^d$ denotes the so-called cell solution, which satisfies

$$\boldsymbol{\nabla} \cdot (a(y) \boldsymbol{\nabla} \boldsymbol{\chi}(y)) = -\nabla_y a(y) \tag{5}$$

and is defined to have zero average. In [7], error bounds for the difference between the solutions to eq. (2) and eq. (3) are proved under certain regularity assumptions. In particular, we have the following result for periodic problems.

Theorem 1 *Given a fixed final time T, assume that $\mathbf{M}^\varepsilon \in C^1([0, T];\ H^2(\Omega))$ is a classical solution to eq. (2) and that there is a constant K independent of ε such that $\|\boldsymbol{\nabla} \mathbf{M}^\varepsilon(\cdot, t)\|_{L^\infty} \leq K$ for all $t \in [0, T]$. Suppose that $\mathbf{M}_0 \in C^\infty(0, T;\ H^\infty(\Omega))$ is a classical solution to eq. (3) and that the assumptions (A1)-(A3) are satisfied. We*

then have for $0 \leq t \leq T$,

$$\|\mathbf{M}^\varepsilon(\cdot,t) - \mathbf{M}_0(\cdot,t)\|_{L^2} \leq C\varepsilon\,, \tag{6}$$

where the constant C *is independent of* ε *and* t *but depends on* K *and* T.

Note that it is easy to show that $\|\boldsymbol{\nabla}\mathbf{M}^\varepsilon(\cdot,t)\|_{L^2} \leq K$ independent of ε, as is for example shown in [7, Appendix B]. Numerically one can check that the same also holds for $\|\boldsymbol{\nabla}\mathbf{M}^\varepsilon(\cdot,t)\|_{L^\infty}$ for typical example problems.

2.2 Heterogeneous Multiscale Methods for the Landau-Lifshitz equation

Heterogeneous Multiscale Methods are a well-established framework for dealing with multiscale problems with scale separation, that involve fast scale oscillations which make it computationally infeasible to properly resolve the problem throughout the whole domain. First introduced by E and Engquist [9], they have since then been applied to problems from many different areas [10, 11].

The general idea of HMM is to approximate the effective solution to the given problem using a coarse scale macro model that is missing some data and is thus incomplete. It is combined with an accurate micro model resolving the fast oscillations in the problem, coupled to the macro solution via the micro initial data. The micro problem is only solved on a small domain around each discrete macro location to keep the computational cost low. The thereby obtained solution is then averaged and provides the information necessary to complete the macro model [9–11].

Since HMM approximates the effective solution to a multiscale problem rather than resolving the fast scales, some error is introduced. In case of the Landau-Lifshitz problem eq. (2) with a periodic material coefficient, the effective solution corresponding to $\mathbf{M}^\varepsilon$ is $\mathbf{M}_0$ satisfying eq. (3). It hence follows from Theorem 1 that the L^2-error between the HMM solution and $\mathbf{M}^\varepsilon$ in the periodic case is always at least $\mathcal{O}(\varepsilon)$.

There are several different HMM models one could choose for the problem eq. (2). Three possibilities are discussed in [8], flux, field and torque model. All three are based on the same micro model, the full Landau-Lifshitz equation eq. (2) which is solved on a time interval $[0,\eta]$, where $\eta \sim \varepsilon^2$. In [7], it is shown that this is the scale of the fast temporal oscillations in the problem. Hence, the micro model is to find $\mathbf{m}^\varepsilon(x,t)$ for $0 \leq t \leq \eta$ such that

$$\partial_t \mathbf{m}^\varepsilon = -\mathbf{m}^\varepsilon \times \boldsymbol{\nabla}\cdot(a^\varepsilon \boldsymbol{\nabla}\mathbf{m}^\varepsilon) - \alpha\mathbf{m}^\varepsilon \times [\mathbf{m}^\varepsilon \times \boldsymbol{\nabla}\cdot(a^\varepsilon\boldsymbol{\nabla}\mathbf{m}^\varepsilon)]\,, \tag{7a}$$

$$\mathbf{m}^\varepsilon(x,0) = \mathbf{m}_{\text{init}}(x) = \Pi^k\mathbf{M}(\cdot,t_j). \tag{7b}$$

The initial data for the micro problem is based on an interpolation of the current macro state $\mathbf{M}$, here denoted by Π^k, which is explained in more detail in Section 4.1. In [8], it is assumed that eq. (7) holds for $x \in \Omega$ with periodic

boundary conditions to simplify the analysis. In practice, one must only solve eq. (7) for $x \in [-\mu', \mu']^d$ to keep down the computational cost. Here $\mu' \sim \varepsilon$, since this is the scale of the fast spatial oscillations in $\mathbf{m}^\varepsilon$. To do this, we have to add artificial boundary conditions which introduce some error as discussed in Sections 4.2 and 4.3.

The three different macro models considered in [8] have the general structure of eqs. (2) and (3) but involve different unknown quantities which have to be obtained by averaging the corresponding data from the micro model eq. (7). In the field model, we have

$$\partial_t \mathbf{M} = -\mathbf{M} \times \mathbf{H}_{\text{avg}}(x, t;\ \mathbf{M}) - \alpha \mathbf{M} \times [\mathbf{M} \times \mathbf{H}_{\text{avg}}(x, t;\ \mathbf{M})], \tag{8a}$$

$$\mathbf{M}(x, 0) = \mathbf{M}_{\text{init}}, \tag{8b}$$

where $\mathbf{H}_{\text{avg}}(x, t;\ \mathbf{M})$ denotes the unknown quantity. In the periodic case, this quantity approximates $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{m} \mathbf{A}^H)$.

In the flux model, eq. (8a) is replaced by

$$\partial_t \mathbf{M} = -\mathbf{M} \times \boldsymbol{\nabla} \cdot \mathbf{F}_{\text{avg}}(x, t;\ \mathbf{M}) - \alpha \mathbf{M} \times [\mathbf{M} \times \boldsymbol{\nabla} \cdot \mathbf{F}_{\text{avg}}(x, t;\ \mathbf{M})], \tag{9}$$

and in the torque model, we instead have

$$\partial_t \mathbf{M} = -\mathbf{T}_{\text{avg}}(x, t;\ \mathbf{M}) - \alpha \mathbf{M} \times \mathbf{T}_{\text{avg}}(x, t;\ \mathbf{M}), \tag{10}$$

where in the periodic case, $\mathbf{F}_{\text{avg}}$ and $\mathbf{T}_{\text{avg}}$ are approximations to $\boldsymbol{\nabla} \mathbf{m} \mathbf{A}^H$ and $\mathbf{m} \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{m} \mathbf{A}^H)$, respectively. As shown in [8], the error introduced when approximating the respective quantities by an averaging procedure, the so-called upscaling error, is bounded rather similarly for all three models, with somewhat lower errors in the flux model. This does not give a strong incentive to choose one of the models over the others. In this paper, we therefore focus on the field model for the following reasons, not related to the upscaling error.

First, when choosing the flux model, the components of the flux should be approximated at different grid locations to reduce the approximation error in the divergence that has to be computed on the macro scale, which typically has a rather coarse discretization. This implies that we need to run separate micro problems for each component of the gradient. This is not necessary when using the field model.

Second, it is seen as an important aspect of micromagnetic algorithms that the norm preservation property of the continuous Landau-Lifshitz equation is mimicked by time integrators for the discretized problem. This is usually achieved by making use of the cross product structure in the equation. However, when choosing the torque model eq. (10), there is no cross product in the first term of the macro model.

The chosen HMM macro model, eq. (8), is discretized on a coarse grid in space with grid spacing ΔX and points $x_i = x_0 + i\Delta X$, where i is a d-dimensional multi-index ranging from 0 to N in each coordinate direction. The

corresponding semi-discrete magnetization values are $\mathbf{M}_i(t) \approx \mathbf{M}(x_i, t)$, which satisfy the semi-discrete equation

$$\partial_t \mathbf{M}_i = -\mathbf{M}_i \times \mathbf{H}_{\text{avg}}(x_i, t;\ \bar{\mathbf{M}}) - \alpha \mathbf{M}_i \times \left[\mathbf{M}_i \times \mathbf{H}_{\text{avg}}(x_i, t;\ \bar{\mathbf{M}})\right], \tag{11a}$$

$$\mathbf{M}_i(0) = \mathbf{M}_{\text{init}}(x_i), \tag{11b}$$

where $\bar{\mathbf{M}}$ denotes the vector containing all the $\mathbf{M}_i$, $i \in \{0, ..., N\}^d$. The notation $\mathbf{H}_{\text{avg}}(x_i, t;\ \bar{\mathbf{M}})$ represents the dependence of $\mathbf{H}_{\text{avg}}$ at location x_i on several values of the discrete magnetization at time t. To discretize eq. (11) in time, we introduce $t_j = t_0 + j\Delta t$, for $j = 0, ..., M$. The specific form of time discretization of eq. (11) is discussed in Section 3.

2.3 Upscaling

To approximate the unknown quantity $\mathbf{H}_{\text{avg}}$ in eq. (11) in an efficient way and to control how fast the approximation converges to the corresponding effective quantity, we use averaging involving kernels as introduced in [12, 13].

Definition 1 ([8, 13]) A function K is in the space of smoothing kernels $\mathbb{K}^{p,q}$ if

1. $K \in C_c^q([-1, 1])$ and $K^{(q+1)} \in BV(\mathbb{R})$.
2. K has p vanishing moments,

$$\int_{-1}^{1} K(x) x^r dx = \begin{cases} 1\,, & r = 0\,, \\ 0\,, & 1 \le r \le p\,. \end{cases}$$

If additionally $K(x) = 0$ for $x \le 0$, then $K \in \mathbb{K}_0^{p,q}$.

We use the conventions that K_μ denotes a scaled version of the kernel K,

$$K_\mu(x) := 1/\mu K(x/\mu),$$

and that in space dimensions with $d > 1$,

$$K(x) := K(x_1) \cdots K(x_d).$$

For the given problem, we choose a kernel $K \in \mathbb{K}^{p_x,q_x}$ for the spatial and $K^0 \in \mathbb{K}_0^{p_t,q_t}$ for the temporal averaging due to the fact that eq. (7) cannot be solved backward in time. The particular upscaling procedure at time t_j is then given by

$$\mathbf{H}_{\text{avg}}(x_i, t_j;\ \bar{\mathbf{M}}) = \int_0^\eta \int_{\Omega_\mu} K_\mu(x) K_\eta^0(t) \boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla} \mathbf{m}^\varepsilon) dx dt, \tag{12}$$

where $\Omega_\mu := [-\mu, \mu]^d$ for a parameter $\mu \sim \varepsilon$ such that $\mu \leq \mu'$, the averaging domain is a subset of the domain that the micro problem is solved on. The micro solution $\mathbf{m}^\varepsilon$ is obtained solving eq. (7) on $[-\mu', \mu']^d \times [0, \eta]$, with initial data $\mathbf{m}_{\text{init}}$ based on $\bar{\mathbf{M}}$ at time t_j and around the discrete location x_i. A second order central difference scheme in space is usually sufficient to obtain an approximation to $\mathbf{m}^\varepsilon$ with errors that are low compared to the averaging errors at a relatively low computational cost.

Assuming instead that the micro problem is solved on $\Omega \times [0, \eta]$, we have the following estimate for the upscaling error for the case of a periodic material coefficient that is proved in [8].

Theorem 2 *Assume that (A1)-(A2) hold and that the micro initial data* $\mathbf{m}_{\text{init}}$ *is normalized such that it satisfies the condition (A3). Let* $\varepsilon^2 < \eta \leq \varepsilon^{3/2}$ *and suppose that for* $x \in \Omega$ *and* $0 \leq t \leq \eta$, *the exact solution to the micro problem eq. (7) is* $\mathbf{m}^\varepsilon(x,t) \in C^1([0,\eta];\ H^2(\Omega))$ *and that there is a constant* c *independent of* ε *such that* $\|\boldsymbol{\nabla}\mathbf{m}^\varepsilon(\cdot,t)\|_{L^\infty} \leq c$. *The solution to the corresponding homogenized problem is* $\mathbf{m}_0 \in C^\infty(0,\eta,H^\infty(\Omega))$. *Moreover, consider averaging kernels* $K \in \mathbb{K}^{p_x,q_x}$ *and* $K^0 \in \mathbb{K}_0^{p_t,q_t}$ *and let* $\varepsilon < \mu < 1$. *Then*

$$\left|\mathbf{H}_{\text{avg}} - \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_{\text{init}}(0)\mathbf{A}^H)\right| =: E_\varepsilon + E_\mu + E_\eta,$$

where the ε*-dependent error* E_ε *and the averaging-domain size dependent error terms* E_μ *and* E_η *are bounded as*

$$E_\varepsilon \leq C\varepsilon, \quad E_\mu \leq C\left(\mu^{p_x+1} + \left(\frac{\varepsilon}{\mu}\right)^{q_x+2}\right) \text{ and } E_\eta \leq C\left(\eta^{p_t+1} + \frac{1}{\mu}\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}\right). \tag{13}$$

In all cases, the constant C *is independent of* ε, μ *and* η *but might depend on* K, K^0 *and* α.

As discussed in [8], for periodic problems we in practice often observe $E_\varepsilon = \mathcal{O}(\varepsilon^2)$ rather than the more pessimistic estimate in the theorem.

Two things are important to note here. First, Theorem 2 states that $\mathbf{H}_{\text{avg}}(x_i, t_j;\ \bar{\mathbf{M}})$ approximates the solution to the corresponding effective quantity involving the micro scale initial data $\mathbf{m}_{\text{init}}$, not the exact macro solution $\mathbf{M}$. We therefore have to require that

$$\partial_x^\beta \mathbf{m}_{\text{init}}(0) = \partial_x^\beta \mathbf{M}(x_i, t_j) \tag{14}$$

for a multi-index β with $|\beta| = 2$ to get an estimate for the actual upscaling error. Bearing in mind a somewhat more general scenario with non-periodic material coefficient where $\mathbf{A}^H$ no longer is constant, we subsequently require eq. (14) to hold for $|\beta| \leq 2$.

Moreover, the quantity that $\mathbf{H}_{\text{avg}}$ in Theorem 2 approximates is independent of α. We can thus choose a different damping parameter in the micro problem than in the macro one to optimize the constants in eq. (13). Typically, it is favorable to have higher damping in the micro problem eq. (7), as is

discussed in the following sections. This can be seen as an introduction of artificial damping to improve numerical properties as is common in for example hyperbolic problems.

2.4 Example problems

Throughout this article, we use three different periodic example problems to illustrate the behavior of the different HMM components and numerical methods under discussion, one 1D example and two 2D examples. Further, non-periodic examples are discussed in Section 5.

(EX1) For the 1D example, the initial data is chosen to be

$$\mathbf{M}_{\text{init}}(x) = \tilde{\mathbf{M}}(x)/|\tilde{\mathbf{M}}(x)|, \quad \tilde{\mathbf{M}}(x) = \begin{bmatrix} 0.5 + \exp(-0.1\cos(2\pi(x-0.32))) \\ 0.5 + \exp(-0.2\cos(2\pi x)) \\ 0.5 + \exp(-0.1\cos(2\pi(x-0.75))) \end{bmatrix}$$

and the material coefficient we consider is $a^\varepsilon(x) = a(x/\varepsilon)$ where

$$a(x) = 1 + 0.5\sin(2\pi x).$$

The corresponding homogenized coefficient, which is not used in the HMM approach but as a reference solution, is $\mathbf{A}^H = \left(\int_0^1 1/a(x)dx\right)^{-1} \approx 0.866$. In Figure 1 the solution $\mathbf{M}^\varepsilon(x,T)$ at $T = 0.1$ and the corresponding HMM approximation $\mathbf{M}(x,T)$ computed on a grid with $\Delta X = 1/24$ are shown. Note that the HMM approximation agrees very well with $\mathbf{M}^\varepsilon$.

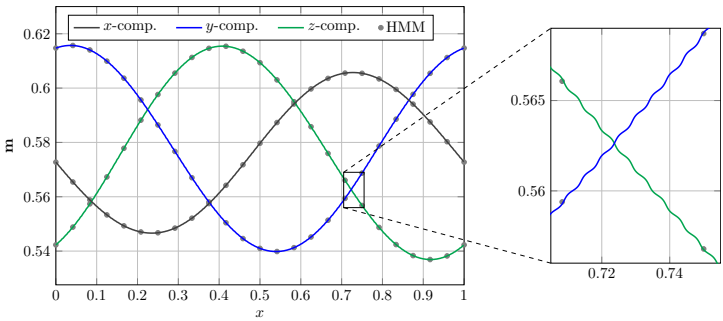


Fig. 1: Solution $\mathbf{m}^\varepsilon$ and corresponding HMM approximation to eq. (1) with setup (EX1) at time $T = 0.1$ when $\varepsilon = 1/200$.

(EX2) For the first 2D example the initial data is

$$\mathbf{m}_{\text{init}}(x) = \tilde{\mathbf{m}}(x)/|\tilde{\mathbf{m}}(x)|,$$

$$\tilde{\mathbf{m}}(x) = \begin{bmatrix} 0.6 + \exp(-0.3(\cos(2\pi(x_1 - 0.25)) + \cos(2\pi(x_2 - 0.12)))) \\ 0.5 + \exp(-0.4(\cos(2\pi x_1) + \cos(2\pi(x_2 - 0.4)))) \\ 0.4 + \exp(-0.2(\cos(2\pi(x_1 - 0.81)) + \cos(2\pi(x_2 - 0.73)))) \end{bmatrix},$$

which is shown in Figure 2. The material coefficient is given by

$$a(x) = 0.5 + (0.5 + 0.25\sin(2\pi x_1))(0.5 + 0.25\sin(2\pi x_2)) + 0.25(\cos(2\pi(x_1 - x_2)) + \sin(2\pi x_1)),$$

which corresponds to a homogenized coefficient with non-zero off-diagonal elements,

$$\mathbf{A}^H \approx \begin{bmatrix} 0.617 & 0.026 \\ 0.026 & 0.715 \end{bmatrix}.$$

For this example, $\mathbf{A}^H$ has to be computed numerically (with high precision).

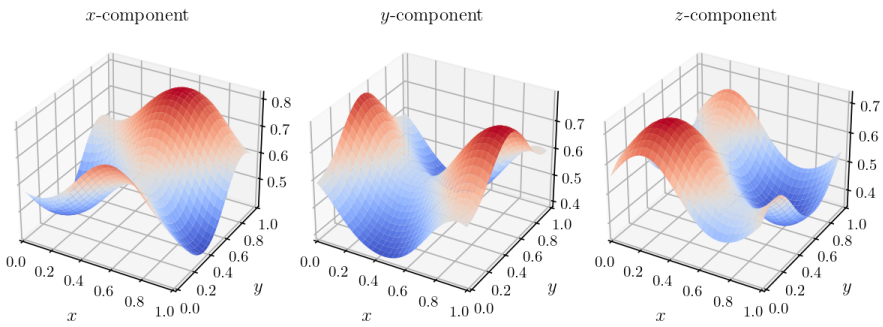


Fig. 2: Initial data $\mathbf{M}_{\text{init}}$ for the 2D problems

(EX3) The second 2D example has the same initial data as (EX2) but a different material coefficient,

$$a(x) = (1.1 + 0.5\sin(2\pi x_1))(1.1 + 0.5\sin(2\pi x_2)).$$

The corresponding homogenized matrix can be computed analytically [14] and takes the value

$$\mathbf{A}^H = 1.1\sqrt{1.1^2 - 0.25}\,\mathbf{I}.$$

In all three cases, it holds that $a^\varepsilon(x) = a(x/\varepsilon)$.

3 Time stepping for Landau-Lifshitz problems

A variety of different methods for time integration of the Landau-Lifshitz equation in a finite difference setting are available, as for example discussed in the review articles [15–17]. Most of these methods can be characterized as either projection methods or geometric integrators, typically based on the

implicit midpoint method. In a projection method, the basic update procedure does not preserve the length of the magnetization vector moments which makes it necessary to project the intermediate result back to the unit sphere at the end of each time step. Commonly used examples for this kind of methods are projection versions of Runge-Kutta methods as well as the Gauss-Seidel projection method [18, 19]. Furthermore, an implicit projection method based on a linear update formula is proposed in [20].

The most common geometric integrator is the implicit midpoint method, which is both norm preserving and in case of no damping, $\alpha = 0$, also energy conserving [21, 22]. However, since it is computationally rather expensive, several semi-implicit variations have been proposed, in particular SIA and SIB introduced in [23] as well as the midpoint extrapolation method, MPE, [24]. Further geometric integrators are the Cayley transform based approaches discussed in [25, 26].

While there are many methods available for time integration of the Landau-Lifshitz equation, which have advantages in different scenarios, we here have a strong focus on computational cost, especially when considering the micro problem, where the subsequent averaging process reduces the importance of conservation of physical properties. For the HMM macro model, the form of the problem, eq. (8), prevents the rewriting of the equation as would be necessary for some integrators, for example the method in [20]. In general, the dependence of $\mathbf{H}_{\text{avg}}$ in eq. (8) on the micro solution makes the use of implicit methods problematic, as is further discussed in Section 3.3.

In the following, we focus on several time integration methods that might be suitable for the given setup and then motivate our choice for the macro and micro problem, respectively.

3.1 Description of selected methods

The methods we focus on are two projection methods, HeunP and RK4P, as well as the semi-implicit midpoint extrapolation method, MPE, introduced in [24] and MPEA, an adaption of the latter method. We furthermore include the implicit midpoint method in the considerations since it can be seen as a reference method for time integration of the Landau-Lifshitz equation.

In this section, we suppose that we work with a discrete grid in space with locations $x_i = x_0 + i\Delta x$, where $i \in \{0, ..., N\}^d$, and consider time points $t_j = t_0 + j\Delta t$, $j = 0, ..., M$. We denote by $\mathbf{m}_i^j \in \mathbb{R}^3$ an approximation to the magnetization at location x_i and time t_j, $\mathbf{m}_i^j \approx \mathbf{m}(x_i, t_j)$. When writing $\mathbf{m}^j$ we refer to a vector in $\mathbb{R}^{3(N+1)^d}$ that contains all $\mathbf{m}_i^j$. In the main part of this section, we do not distinguish between macro and micro problem but focus on the general behavior of the time stepping methods. Thus, $\mathbf{m}$ can denote both a micro or macro solution.

We furthermore use the notation $\mathbf{f}_i(\mathbf{m}^j)$ to denote the value of a function $\mathbf{f}$ at location x_i which might depend on the values of $\mathbf{m}^j$ at several space locations. In particular, we write $\mathbf{H}_i(\mathbf{m}^j)$ to denote a discrete approximation to the effective field at x_i. On the macro scale, it thus holds that $\mathbf{H}_i(\mathbf{m}^j) \approx$

$\mathbf{H}_{\text{avg}}(x_i, t_j;\ \bar{\mathbf{M}})$ or, when considering the corresponding homogenized problem in case of a periodic material coefficient,

$$\mathbf{H}_i(\mathbf{m}^j) \approx \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}(x_i, t_j)\mathbf{A}^H).$$

For the micro problem, we have

$$\mathbf{H}_i(\mathbf{m}^j) \approx \boldsymbol{\nabla} \cdot (a(x_i/\varepsilon)\boldsymbol{\nabla}\mathbf{m}^\varepsilon(x_i, t_j)).$$

The particular form of $\mathbf{H}_i(\mathbf{m}^j)$ does not have a major influence on the following discussions if not explicitly stated otherwise.

HeunP and RK4P

HeunP and RK4P are the standard Runge Kutta 2 and Runge Kutta 4 methods with an additional projection back to the unit sphere at the end of every time step. Let $\mathbf{f}(\mathbf{m}^j)$ be the $3(N+1)^d$-vector such that

$$\mathbf{f}_i(\mathbf{m}^j) := -\mathbf{m}_i^j \times \mathbf{H}_i(\mathbf{m}^j) - \alpha \mathbf{m}_i^j \times \left[\mathbf{m}_i^j \times \mathbf{H}_i(\mathbf{m}^j)\right], \qquad i \in \{0, ..., N\}^d.$$

Then in the Runge-Kutta methods, one computes stage values

$$\mathbf{k}_1 = \mathbf{f}(\mathbf{m}^j), \quad \mathbf{k}_2 = \mathbf{f}(\mathbf{m}^j + \tfrac{\Delta t}{2}\mathbf{k}_1), \quad \mathbf{k}_3 = \mathbf{f}(\mathbf{m}^j + \tfrac{\Delta t}{2}\mathbf{k}_2), \quad \mathbf{k}_4 = \mathbf{f}(\mathbf{m}^j + \Delta t \mathbf{k}_3).$$

In HeunP (RK2P), the time step update then is given by

$$\mathbf{m}_i^{j+1} = \tilde{\mathbf{m}}_i/|\tilde{\mathbf{m}}_i|, \qquad \text{where} \qquad \tilde{\mathbf{m}} = \mathbf{m}^j + \tfrac{\Delta t}{2}(\mathbf{k}_1 + \mathbf{k}_2), \tag{15}$$

and for RK4P,

$$\mathbf{m}_i^{j+1} = \tilde{\mathbf{m}}_i/|\tilde{\mathbf{m}}_i|, \qquad \text{where} \qquad \tilde{\mathbf{m}} = \mathbf{m}^j + \tfrac{\Delta t}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4). \tag{16}$$

HeunP is a second order method and RK4P is fourth order accurate.

Implicit midpoint

Using the implicit midpoint method, the Landau-Lifshitz equation is discretized as

$$\frac{\mathbf{m}_i^{j+1} - \mathbf{m}_i^j}{\Delta t} = -\frac{\mathbf{m}_i^j + \mathbf{m}_i^{j+1}}{2} \times \mathbf{h}_i\left(\frac{\mathbf{m}^j + \mathbf{m}^{j+1}}{2}\right), \tag{17}$$

where

$$\mathbf{h}_i(\mathbf{m}) := \mathbf{H}_i(\mathbf{m}) + \alpha \mathbf{m}_i \times \mathbf{H}_i(\mathbf{m}). \tag{18}$$

Hence, the values for $\mathbf{m}_i^{j+1}$ are obtained by solving the nonlinear system

$$\mathbf{m}^{j+1} = \left(\mathbf{I} + \tfrac{\Delta t}{2}\left[\mathbf{h}\left(\tfrac{\mathbf{m}^j+\mathbf{m}^{j+1}}{2}\right)\right]_\times\right)^{-1}\left(\mathbf{I} - \tfrac{\Delta t}{2}\left[\mathbf{h}\left(\tfrac{\mathbf{m}^j+\mathbf{m}^{j+1}}{2}\right)\right]_\times\right)\mathbf{m}^j, \quad (19)$$

where $[\mathbf{h}]_\times$ is the matrix such that the matrix-vector product $[\mathbf{h}]_\times \mathbf{m}^j$ corresponds to taking the cross products $\mathbf{h}_i \times \mathbf{m}_i^j$ for all $i \in \{0, ..., N\}^d$. The implicit midpoint method is norm-conserving and results in a second order accurate approximation.

However, when using Newton's method to solve the non-linear system eq. (19), one has to compute the Jacobian of the right-hand side with respect to $\mathbf{m}^{j+1}$, a sparse, but not (block)-diagonal, $3(N+1)^d \times 3(N+1)^d$ matrix and then solve the corresponding linear system in each iteration, which has a rather high computational cost. In case of the HMM macro model, the Jacobian cannot be computed analytically since then $\mathbf{H}_i$ in eq. (18) is replaced by the averaged quantity $\mathbf{H}_{\text{avg}}(x_i, t_j;\ \mathbf{m})$, with a dependence on $\mathbf{m}$ that is very complicated. A numerical approximation is highly expensive since it means solving CN^{2d} additional micro problems per time step.

MPE and MPEA

As described in [24], the idea behind the midpoint extrapolation method is to approximate $\mathbf{h}(\frac{\mathbf{m}^j+\mathbf{m}^{j+1}}{2})$ in eq. (17) using the explicit extrapolation formula

$$\mathbf{h}\left(\frac{\mathbf{m}^j + \mathbf{m}^{j+1}}{2}\right) \approx \mathbf{h}^{j+1/2} := \tfrac{3}{2}\mathbf{h}(\mathbf{m}^j) - \tfrac{1}{2}\mathbf{h}(\mathbf{m}^{j-1}). \quad (20)$$

The update for $\mathbf{m}_i$ then becomes

$$\frac{\mathbf{m}_i^{j+1} - \mathbf{m}_i^j}{\Delta t} = -\frac{\mathbf{m}_i^j + \mathbf{m}_i^{j+1}}{2} \times \mathbf{h}_i^{j+1/2}. \quad (21)$$

The quantity $\mathbf{h}_i^{j+1/2}$ here is independent of $\mathbf{m}^{j+1}$, which means that the problem decouples into $(N+1)^d$ small 3×3 systems. Since the first term on the right-hand side still contains $\mathbf{m}_i^{j+1}$, this is considered a semi-implicit method. Just as the implicit midpoint method, MPE is second order accurate. We furthermore propose to use third-order accurate extrapolation,

$$\mathbf{h}^{j+1/2} := \frac{23}{12}\mathbf{h}(\mathbf{m}^j) - \frac{16}{12}\mathbf{h}(\mathbf{m}^{j-1}) + \frac{5}{12}\mathbf{h}(\mathbf{m}^{j-2}), \quad (22)$$

in eq. (21), which gives MPEA, the adapted MPE method. As it is based on the implicit midpoint method, MPEA is second order accurate just like MPE, but has better stability properties for low damping, as shown in the next section.

As MPE and MPEA are multi-step methods, values for $\mathbf{m}^1$ (and $\mathbf{m}^2$) are required for startup. These can be obtained using HeunP or RK4P, as suggested in [24].

Comparison of methods

In Figure 3, the error with respect to a reference solution $\mathbf{m}_{\text{ref}}$ is shown for all the considered (semi-)explicit methods and two example problems, a 1D problem and a 2D problem. Both example problems are homogenized problems on a rather coarse spatial discretization grid such as we might have in the HMM macro problem. For the 1D problem, the implicit midpoint method is also included for reference. For small time step sizes, one can observe the expected

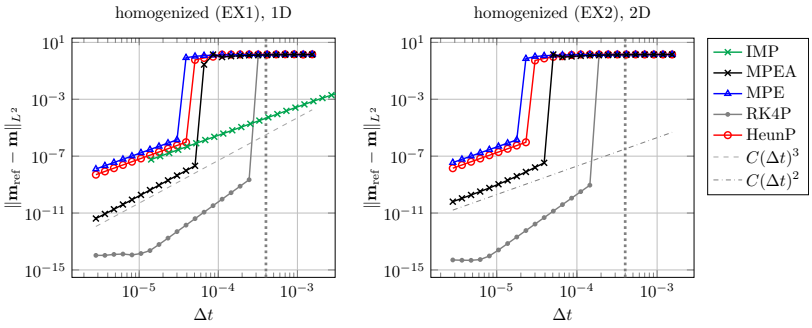


Fig. 3: Comparison of L^2-error in different time stepping methods when varying the time step size Δt given a fixed spatial discretization with $\Delta x = 1/50$ and damping parameter $\alpha = 0.01$. The dotted vertical line corresponds to $\Delta t = (\Delta x)^2$ in each case.

convergence rates for HeunP, RK4P and MPE. For MPEA, we observe third order convergence in the 1D problem, while in the 2D case, we have second order convergence for small Δt. Overall, MPEA results in lower errors compared to HeunP, MPE and the implicit midpoint method. RK4P is most accurate.

3.2 Stability of the time stepping methods

It is well known that in explicit time stepping methods for the Landau-Lifshitz equation, the choice of time step size Δt given Δx in space is severely constrained by numerical stability, see for example [17, 18]. Note that due to the norm preservation property of the considered methods, the solutions cannot grow arbitrarily as unstable solutions often do in other applications. However, when taking too large time steps, explicit time integration will typically result in solutions that oscillate rapidly and do not represent the intended solution in any way. Following standard practice in the field, we refer to this behavior

as instability in this section. This stability limit is seen clearly for all methods expect IMP in Figure 3.

In numerical experiments, we observe that in order to obtain stable solutions, the time step size Δt has to be chosen proportional to Δx^2 for all of the considered methods, both explicit and semi-explicit. This is exemplified in Figure 4 for (EX1) with $\alpha = 0.01$. To get a better understanding of stability,

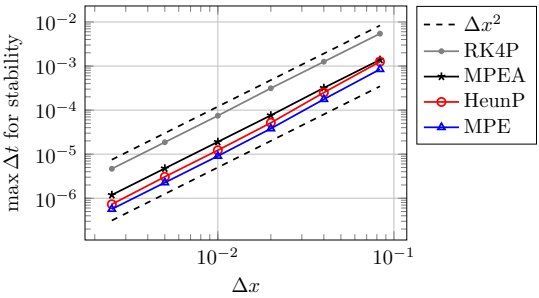


Fig. 4: Empirically found maximum value of Δt that still results in a stable solution for varying values of Δx in (EX1), homogenized, with $\alpha = 0.01$.

consider a semi-discrete form of the Landau-Lifshitz equation in one dimension with a constant material coefficient, equal to one. This can be written as

$$\partial_t \mathbf{m} = \mathbf{f}_\alpha(\mathbf{m}), \qquad \mathbf{f}_\alpha(\mathbf{m}) = -\mathbf{m}\times D\mathbf{m} - \alpha\mathbf{m}\times[\mathbf{m}\times D\mathbf{m}] = -B(\mathbf{m})D\mathbf{m}, \quad (23)$$

where $\mathbf{m} \in \mathbb{R}^{3N}$ contains the vectors $\{\mathbf{m}_i\}$, D is the discrete Laplacian and $B(\mathbf{m})$ is a block diagonal skew-symmetric matrix with eigenvalues $\{0,\ +i-\alpha,\ -i-\alpha\}$ that comes from the cross products. If $\{\mathbf{m}_i\}$ samples a smooth function, the Jacobian of $\mathbf{f}_\alpha(\mathbf{m})$ can be approximated as

$$\nabla_{\mathbf{m}}\mathbf{f}_\alpha(\mathbf{m}) = -\nabla_{\mathbf{m}}B(\mathbf{m})D\mathbf{m} \approx -B(\mathbf{m})D.$$

Still assuming smoothness, one can subsequently deduce [27] that the eigenvalues of the Jacobian are approximately given as the eigenvalues ω of $-D$ multiplied by the eigenvalues of $B(\mathbf{m})$, namely

$$\lambda_+ \approx (i-\alpha)\omega, \qquad \lambda_- \approx (-i-\alpha)\omega, \qquad \lambda_0 \approx 0.$$

The eigenvalues of $-D$, the negative discrete Laplacian, are real, positive and bounded by $O(\Delta x^{-2})$. Consequently, the eigenvalues of the Jacobian $\nabla_{\mathbf{m}}\mathbf{f}_\alpha(\mathbf{m})$ will lie along the lines $s(\pm i-\alpha)$ for real $s \in [0, O(\Delta x^{-2})]$ in the complex plane. This is illustrated in Figure 5a where we have plotted the eigenvalues of $\nabla_{\mathbf{m}}\mathbf{f}_\alpha(\mathbf{m})$, scaled by Δx^2, for several values of α. One can observe that given $\alpha = 0$, the eigenvalues are purely imaginary. As α increases, the real parts of the eigenvalues decrease correspondingly.

For the Landau-Lifshitz equation eq. (2) with a material coefficient as well as the homogenized equation eq. (3), the eigenvalues of the corresponding Jacobians get a different scaling based on the material coefficient but their general behavior is not affected. We hence conjecture that it is necessary that

$$\frac{\Delta t}{(\Delta x)^2} \leq C_{\mathrm{stab},\alpha}, \tag{24}$$

where $C_{\mathrm{stab},\alpha}$ is a constant depending on the chosen integrator, the damping parameter α and the material coefficient. Based on several numerical examples, we observe for the latter dependence that

$$C_{\mathrm{stab},\alpha} \lesssim C_\alpha \begin{cases} (\max_{y\in Y}|a(y)|)^{-1}, & \text{original problem,} \\ (\max_{i,j}|\mathbf{A}^H_{ij}|)^{-1}, & \text{homogenized problem,} \end{cases}$$

where C_α denotes further dependence on α and the integrator.

Stability regions of related methods

In order to better understand the stability behavior of the considered time integrators, it is beneficial to study the stability regions of some well-known, related methods. For HeunP and RK4P, we regard the corresponding integrators without projection. We observe a very similar stability behavior when using Heun and RK4 to solve the problems considered in Figures 3 and 4.

To get some intuition about MPE and MPEA, we start by considering the problem

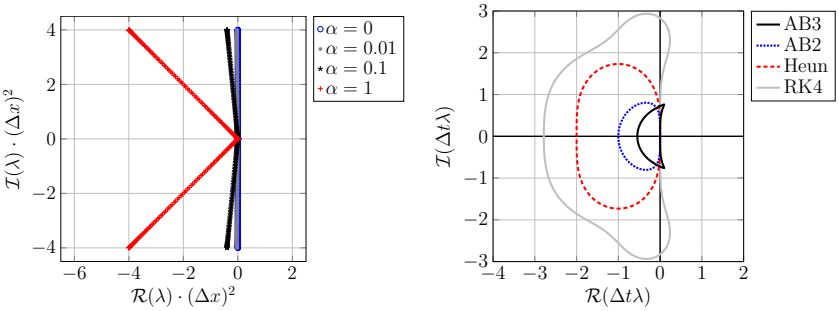
$$\partial_t \mathbf{m} = -\mathbf{n} \times \mathbf{H}(\mathbf{m}) - \alpha \mathbf{n} \times [\mathbf{m} \times \mathbf{H}(\mathbf{m})] = -\mathbf{n} \times \mathbf{h}(\mathbf{m}), \tag{25}$$

where $\mathbf{n}$ is a given vector function, constant in time, with $|\mathbf{n}| = 1$. This corresponds to replacing the first $\mathbf{m}$ in each term on the right-hand side in eq. (2) by a constant approximation. For this problem, time stepping according to the MPE update eq. (21) results in

$$\frac{\mathbf{m}_i^{j+1} - \mathbf{m}_i^j}{\Delta t} = -\mathbf{n}_i \times \mathbf{h}_i^{j+1/2} = -\frac{3}{2}\left(\mathbf{n}_i \times \mathbf{h}_i(\mathbf{m}^j)\right) + \frac{1}{2}\left(\mathbf{n}_i \times \mathbf{h}_i(\mathbf{m}^{j-1})\right). \tag{26}$$

This is the same update scheme as one gets when applying the Adams-Bashforth 2 (AB2) method to eq. (25). In the same way, MPEA and AB3 are connected. Furthermore, note that the term that was replaced by $\mathbf{n}$ in eq. (2) to obtain eq. (25) is the one that is treated implicitly in the semi-implicit methods MPE and MPEA. We hence expect that studying the stability of AB2 and AB3 can give an indication of what to expect for MPE and MPEA. This is backed up by the fact that the stability properties observed for MPE and MPEA in Figures 3 and 4 are closely matched when using AB2 and AB3, respectively, to solve the respective problems.

When comparing the stability regions of RK4, Heun, AB2 and AB3 as shown in Figure 5b, one clearly sees that the Runge-Kutta methods have larger stability regions than the multi-step methods. RK4's stability region is largest and contains part of the imaginary axis, while the one for Heun is only close to the imaginary axis in a shorter interval. AB2 and AB3 have stability regions



(a) Eigenvalues of the Jacobian $\nabla \mathbf{f}_\alpha(\mathbf{m})$ for several values of α

(b) Stability regions of RK2 (Heun), RK4 as well as AB2 and AB3.

Fig. 5: Eigenvalues of Jacobian $\nabla \mathbf{f}_\alpha(\mathbf{m})$ and stability regions of methods related to the considered time integrators.

with a similar extent in the imaginary direction, but while AB3's contains part of the imaginary axis, AB2's does not. On the other hand, the stability region of AB2 is wider in the real direction.

Consider now again the example problems shown in Figure 3 where $\alpha = 0.01$, which implies that the eigenvalues of the corresponding Jacobian are rather close to the imaginary axis. Based on Figure 5b we therefore expect the methods with related stability areas which include parts of the imaginary axis, RK4P and MPEA, to require fewer time steps than HeunP and MPE. This matches with the observed stability behavior in Figure 3.

Influence of α

To further investigate the influence of α on $C_{\text{stab},\alpha}$, this factor is shown in Figure 6 for varying α, both for the considered methods HeunP, RK4P, MPE and MPEA and the discussed related methods, Heun, RK4, AB2 and AB3. The behavior of $C_{\text{stab},\alpha}$ is almost the same for the actual and the related methods.

As expected, we observe that $C_{\text{stab},\alpha}$ for low α-values is constant for MPEA and RK4P, with related stability regions that include the imaginary axis, while for HeunP and MPE, lower α results in lower $C_{\text{stab},\alpha}$. When increasing α, the eigenvalues of $\boldsymbol{\nabla}_m \mathbf{f}_\alpha$ as defined in eq. (23) get larger real parts and the stability regions' extent in the real direction becomes more important. For RK4P and MPEA, this means that for $\alpha \gtrsim 0.2$, $C_{\text{stab},\alpha}$ decreases as α increases. For HeunP and MPE, the highest $C_{\text{stab},\alpha}$ is obtained at $\alpha \approx 0.5$. For higher α, the

required $C_{\text{stab},\alpha}$ decreases as α increases. Overall, MPEA requires the lowest $C_{\text{stab},\alpha}$ for high α, which agrees with the fact that the related stability region is shortest in the real direction. The highest $C_{\text{stab},\alpha}$ is still the one for RK4P, in accordance with the stability region considerations.

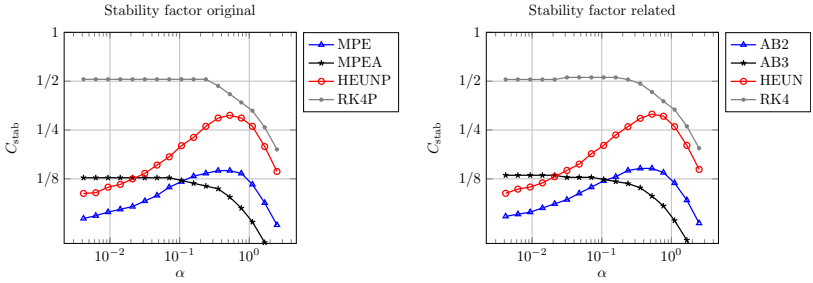


Fig. 6: Dependence of $C_{\text{stab},\alpha}$ on α, for actual and related time integrators. Based on the homogenized solution to (EX2).

However, HeunP is a two-stage method and each RK4P step consists of four stages, while MPE and MPEA are multi-step methods that only require computation of one new stage per time step. In general, each step of HeunP/RK4P thus has roughly two/four times the computational cost as a MPE or MPEA step. To take this into account, we compare the total number of computations for each method by considering the factor $s/C_{\text{stab},\alpha}$, where s denotes the number of stages in the method. This is shown in Figure 7.

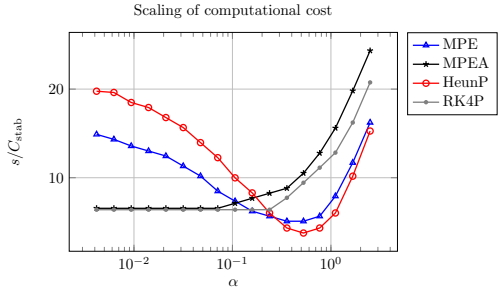


Fig. 7: Left: Dependence of $C_{\text{stab},\alpha}$ on α, based on the homogenized solution to (EX2). Right: corresponding scaling of computational cost.

We hence draw the following conclusion.

- For $\alpha < 0.1$, RK4 and MPEA result in approximately the same computational cost, independent of the specific value of α, while MPE and HeunP require significantly more computations.

- For high (artificial) damping, the situation changes and HeunP has the lowest computational cost of the considered time integrators.

3.3 Macro time stepping

On the HMM macro scale, the given spatial discretization is in general rather coarse, containing only relatively few grid points, and we are interested in longer final times. Therefore, an implicit method such as IMP might seem suitable here. However, the resulting computational cost is higher than with an explicit method since we cannot compute the required Jacobian analytically as discussed in Section 3.1.

When considering which (semi)-explicit method is most suitable, we have to consider the value of the damping constant α. According to for example [28, 29], the value of α is less than 0.1 or even 0.01 for typical metallic materials such as Fe, Co and Ni, and could be one or two orders smaller for ferromagnetic oxides or garnets. On the macro scale, we hence typically reckon with α between 10^{-1} and 10^{-4}. This also matches with the α-values typically used in the literature, see for instance [15, 18, 30] and [24]. For this range of α, we conclude based on the discussion in Section 3.2 that RK4P and MPEA are preferred for time integration on the macro scale. These are also the methods which give the most accurate solutions as, for example, shown in Figure 3.

For the overall error on the macro scale in the periodic case, we expect that

$$\begin{aligned} \|\mathbf{M}^{\varepsilon} - \mathbf{M}_0\| &\leq C\left(\varepsilon + (\Delta t)^k + (\Delta X)^{\ell} + e_{\mathrm{HMM}}\right) \\ &\leq C\left(\varepsilon + (\Delta X)^{\min(2k,\ell)} + e_{\mathrm{HMM}}\right), \end{aligned} \tag{27}$$

where the factor ε follows from Theorem 1 and k is the order of accuracy of the time integrator. Moreover, ℓ is the order of accuracy of the spatial approximation to the effective field on the macro scale and e_{HMM} is an additional error due to the fact that we approximate this effective field by $\mathbf{H}_{\mathrm{avg}}$ in the upscaling procedure.

Because of the time step restriction required for stability in explicit and semi-explicit methods, eq. (24), it is desirable to have relatively large Δx to reduce computational cost. We therefore propose to use a higher order method in space, ideally $\ell = 2k$.

We hence have two possible choices. One can either use the fourth order accurate RK4P in combination with an eighth order approximation in space to get a macro scheme with very high order of accuracy regarding space and time discretization. However, to get the full effect of this, $\mathbf{H}_{\mathrm{avg}}$ has to be approximated very precisely such that also e_{HMM} is low, see also Section 4.5, which in turn can result in rather high computational cost. Alternatively, one can apply the second order MPEA and a spatial approximation such that $\ell = 4$ which results in fourth order accuracy of the space and time discretization. Since MPEA also has the advantage that it is a geometric integrator and norm preserving without projection, this is what we propose to use.

3.4 Micro time stepping

When considering the micro problem, there are two important differences compared to the macro problem. First, the fast oscillations in the solution $\mathbf{m}^\varepsilon$ are on different scales in time and space. The time scale we are interested in is $\mathcal{O}(\varepsilon^2)$ while the spatial scale is $\mathcal{O}(\varepsilon)$. Therefore a time step size proportional to $(\Delta x)^2$ is suitable to obtain a proper resolution of the fast oscillations in time. Second, as discussed Section 2.3, we can choose the damping parameter α for the micro problem to optimize convergence of the upscaling errors. As shown in Section 4, it is typically advantageous to use artificial damping and set α close to one, considerably higher than in the macro problem.

The order of accuracy of the time integrator is not an important factor, since already for a second order accurate integrator, the time integration error usually is significantly lower than the space discretization error due to the given time step restriction. As the micro problem is posed on a relatively short time interval and the solution then is averaged in the upscaling process, inherent norm preservation that geometric integrators have is not an important factor here either. The considerations in Section 3.2 thus imply that the optimal strategy with respect to computational cost is to use HeunP for time integration when $\alpha > 0.2$ is chosen in the micro problem.

4 Micro problem setup

In this section, we investigate how different aspects of the micro problem influence the upscaling error as well as the overall macro solution. In particular, the choice of initial data and the size of the computational and averaging domain in space and time are important. Consider the periodic case, $a^\varepsilon(x) = a(x/\varepsilon)$. Then it holds for the error in the HMM approximation to the effective field that

$$\begin{aligned} E_{\text{approx}} &:= \left|\mathbf{H}_{\text{avg}} - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{M}_0(x_i,t_j)\mathbf{A}^H)\right| \\ &\leq \left|\mathbf{H}_{\text{avg}} - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_{\text{init}}(0)\mathbf{A}^H)\right| \\ &\quad + \left|\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_{\text{init}}(0)\mathbf{A}^H) - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{M}_0(x_i,t_j)\mathbf{A}^H)\right| \\ &=: E_{\text{avg}} + E_{\text{disc}}. \end{aligned} \tag{28}$$

The discretization error E_{disc} is determined by the choice of initial data $\mathbf{m}_{\text{init}}$ to the micro problem and is analyzed in the next section, Section 4.1. Given that we have initial data with $|\mathbf{m}_{\text{init}}| = 1$ and an exact solution to the micro problem on the whole domain, the averaging error E_{avg} can be bounded using Theorem 2. When solving the micro problem numerically and only on a subdomain $[-\mu', \mu']^d$, additional errors are introduced. We can split E_{avg} as

$$E_{\text{avg}} = E_\varepsilon + E_\mu + E_\eta + E_{\mu'} + E_{\text{num}}, \tag{29}$$

where E_ε, E_μ and E_η are as in Theorem 2. They depend on ε and the parameters μ and η, which determine the size of the micro problem averaging domains in space and time. Moreover, the choice of averaging kernels, K and K^0, influences these errors. How to specifically choose these parameters is discussed in Section 4.3. The term $E_{\mu'}$ comprises errors due to the micro problem boundary conditions, as explained in Section 4.2, and E_{num} errors due to the numerical discretization of the micro problem, which is done using a standard second order finite difference approximation to $\boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla} \mathbf{m}^\varepsilon)$ and HeunP for time integration.

4.1 Initial data

We first consider how to choose the initial data $\mathbf{m}_{\text{init}}$ to a micro problem based on the current macro state, obtained according to eq. (11). We here suppose that the current given discrete magnetization values match with the (exact) macro solution at time $t = t_j$, $\mathbf{M}_i = \mathbf{M}(x_i, t_j)$. For the discretization of $\mathbf{M}$, the same setup as described in Section 2.2 for the HMM macro model is considered.

The initial data $\mathbf{m}_{\text{init}}$ for the micro problem should satisfy two conditions.

1. It should be normalized, $|\mathbf{m}_{\text{init}}(x)| = 1$ for all $x \in [-\mu', \mu']^d$, to satisfy the conditions necessary for Theorem 2, which we use to bound E_{avg}.
2. The initial data should be consistent with the current macro solution in the sense that given a multi-index β with $|\beta| \leq 2$,

$$\left|\partial_x^\beta \mathbf{m}_{\text{init}}(0) - \partial_x^\beta \mathbf{M}(x_k, t_j)\right| = \mathcal{O}((\Delta X)^\ell). \tag{30}$$

   Then the discretization error in eq. (28) is $E_{\text{disc}} = \mathcal{O}((\Delta X)^\ell)$. As described in eq. (27) in Section 3.3, when using a kth order explicit time stepping method, it is ideal in terms of order of accuracy to choose $\ell = 2k$.

In order to get initial data satisfying the requirements, an approach based on polynomial interpolation is applied. We use $p^{[n]}(x)$ to denote an interpolating polynomial of order n, and let $\mathbf{P}_{[n]}(x) = \left[p_1^{[n]}, p_2^{[n]}, p_3^{[n]}\right]^T$, a vector containing an independent polynomial for each component in $\mathbf{M}$, such that

$$\mathbf{P}_{[n]}(x_i) = \mathbf{M}_i, \quad i = 0, \ldots, n.$$

When $d > 1$, we apply one-dimensional interpolation in one space dimension after the other. For matters of simplicity, we regard a 1D problem in the following analytical error estimates. Due to the tensor product extension, the considerations generalize directly to higher dimensions.

Without loss of generality, we henceforth assume that we want to find initial data for the micro problem associated with the macro grid point at location x_k based on $2k$-th order polynomial interpolation. This implies that the macro grid points involved in the process are x_j, $j = 0, ..., 2k$.

According to standard theory, it holds for the interpolation errors that given $0 \leq i \leq 2k$ and $\mathbf{M} \in C^{(2k+1)}([x_0, x_{2k}])$,

$$\sup_{x \in [x_0, x_{2k}]} |\mathbf{P}^{(i)}_{[2k]}(x) - \mathbf{M}^{(i)}(x, t_j)| \leq C(\Delta X)^{2k+1-i}. \tag{31}$$

Furthermore, it is well known that given a $2k$-th degree interpolating polynomial, it holds that

$$\mathbf{P}'_{[2k]}(x_k) = D_{[2k]}\mathbf{M}(x_k, t_j), \qquad \mathbf{P}''_{[2k]}(x_k) = D^2_{[2k]}\mathbf{M}(x_k, t_j),$$

where $D_{[2k]}$ and $D^2_{[2k]}$ denote the $2k$-th order standard central finite difference approximations to the first and second derivative, see for example [31]. As a direct consequence, we have in the grid point x_k,

$$|\mathbf{M}'(x_k, t_j) - \mathbf{P}'_{[2k]}(x_k)| \leq C(\Delta X)^{2k}, \qquad |\mathbf{M}''(x_k, t_j) - \mathbf{P}''_{[2k]}(x_k)| \leq C(\Delta X)^{2k}. \tag{32}$$

Note that this gives a better bound for the error in the second derivative in the point x_k than eq. (31), valid on the whole interval $[x_0, x_{2k}]$. The bounds in Equation (32) show that we have the required consistency, eq. (30), between macro and micro derivatives when directly using $\mathbf{P}_k(x)$ to obtain the initial data for the micro problem. However, the disadvantage of this approach is that the polynomial vector $\mathbf{P}_k$ is not normalized. For the deviation of its length from one, it holds by eq. (31) that

$$\big||\mathbf{P}_{[2k]}| - 1\big| = \big||\mathbf{P}_{[2k]}| - |\mathbf{M}|\big| \leq |\mathbf{P}_{[2k]} - \mathbf{M}| \leq C(\Delta X)^{2k+1}.$$

Consider instead a normalized function $\mathbf{Y}(x)$ for which $|\mathbf{Y}(x)| = 1$. Then the derivative $\mathbf{Y}'$ is orthogonal to $\mathbf{Y}$ as it holds that

$$\mathbf{Y}'(x) \cdot \mathbf{Y}(x) = \frac{1}{2}\frac{d}{dx}|\mathbf{Y}(x)|^2 = 0.$$

In particular, this shows that $\mathbf{M}' \cdot \mathbf{M} = 0$. However, in general,

$$\mathbf{M}(x_k, t_j) \cdot D_{[2k]}\mathbf{M}(x_k, t_j) = \mathbf{P}_{[2k]}(x_k) \cdot \mathbf{P}'_{[2k]}(x_k) \neq 0.$$

Hence there is no normalized interpolating function $\mathbf{Y}$ such that $\mathbf{Y}'(x_{\mathrm{k}})$ becomes a standard linear $2k$-th order central difference approximation, $D_{[2k]}$. In the following, we consider the normalized interpolating function $\mathbf{Q}_{[n]}(x)$ defined as

$$\mathbf{Q}_{[n]}(x) := \mathbf{P}_{[n]}(x)/|\mathbf{P}_{[n]}(x)|\,, \tag{33}$$

and show that it satisfies the consistency requirement eq. (30).

Lemma 1 In one space dimension, the normalized function $\mathbf{Q}_{[2k]}$ satisfies eq. (30) with $\ell = 2k$.

Proof As $\left|\mathbf{P}_{[2k]}(x_i)\right| = |\mathbf{M}_i| = 1$ in the grid points x_i, where $i = 0, ..., 2k$, it follows directly that $\mathbf{Q}_{[2k]}(x_i) = \mathbf{M}_i$ for $i = 0, ..., 2k$, the normalized function still interpolates the given points.

For the first derivative of $\mathbf{Q}_{[2k]}$, it holds that

$$\mathbf{Q}'_{[2k]} = \frac{d}{dx}\left(\mathbf{P}_{[2k]}/|\mathbf{P}_{[2k]}|\right) = \mathbf{P}'_{[2k]}/|\mathbf{P}_{[2k]}| - (\mathbf{P}^T_{[2k]}\mathbf{P}'_{[2k]})\mathbf{P}_{[2k]}/|\mathbf{P}_{[2k]}|^3,$$

which, together with the fact that $\mathbf{P}_{[2k]}(x_i) = \mathbf{M}_i$ for $i = 0, .., 2k$, implies that

$$\mathbf{Q}'_{[2k]}(x_i) = (\mathbf{I} - \mathbf{M}_i\mathbf{M}_i^T)\mathbf{P}'_{[2k]}(x_i). \tag{34}$$

In particular, it holds due to orthogonality that

$$\begin{aligned}\left|\mathbf{Q}'_{[2k]}(x_i) - \mathbf{M}'(x_i)\right| &= \left|(\mathbf{I} - \mathbf{M}_i\mathbf{M}_i^T)\left(\mathbf{P}'_{[2k]}(x_i) - \mathbf{M}'(x_i)\right)\right| \\ &\leq C|\mathbf{P}'_{[2k]}(x_i) - \mathbf{M}'(x_i)| \leq C(\Delta X)^{2k},\end{aligned}$$

hence $\mathbf{Q}'_{[2k]}(x_k)$ is a $2k$-th order approximation to the derivative of $\mathbf{M}$ in x_k. For the second derivative of $\mathbf{Q}_{[n]}$, it holds in general that

$$\mathbf{Q}''_{[n]} = \frac{\mathbf{P}''_{[n]}}{|\mathbf{P}_{[n]}|} - 2\frac{(\mathbf{P}_{[n]}\cdot\mathbf{P}'_{[n]})\mathbf{P}'_{[n]}}{|\mathbf{P}_{[n]}|^3} + 3\frac{(\mathbf{P}_{[n]}\cdot\mathbf{P}'_{[n]})^2\mathbf{P}_{[n]}}{|\mathbf{P}_{[n]}|^5} - \frac{(|\mathbf{P}'_{[n]}|^2 + \mathbf{P}''_{[n]}\cdot\mathbf{P}_{[n]})\mathbf{P}_{[n]}}{|\mathbf{P}_{[n]}|^3}, \tag{35}$$

where we can rewrite

$$|\mathbf{P}'_{[n]}|^2 + \mathbf{P}''_{[n]}\cdot\mathbf{P}_{[n]} = (\mathbf{P}_{[n]} - \mathbf{M})\cdot\mathbf{P}''_{[n]} + (\mathbf{P}'_{[n]} - \mathbf{M}')\cdot(\mathbf{M}' + \mathbf{P}'_{[n]}) + (\mathbf{P}''_{[n]} - \mathbf{M}'')\cdot\mathbf{M},$$

For a $2k$-th order interpolating polynomial, we hence have in the grid point x_k that

$$\begin{aligned}&\left|\left(|\mathbf{P}'_{[2k]}(x_k)|^2 + \mathbf{P}''_{[2k]}(x_k)\cdot\mathbf{P}_{[2k]}(x_k)\right)\mathbf{P}_{2k}(x_k)\right| \\ &\quad\leq \left|(\mathbf{P}'_{[2k]}(x_k) - \mathbf{M}'(x_k))\cdot(\mathbf{M}'(x_k) + \mathbf{P}'_{[2k]}(x_k)) + (\mathbf{P}''_{[2k]}(x_k) - \mathbf{M}''(x_k))\cdot\mathbf{M}(x_k)\right| \\ &\quad\leq C\left(\left|(\mathbf{P}'_{[2k]}(x_k) - \mathbf{M}'(x_k))\right| + \left|\mathbf{P}''_{[2k]}(x_k) - \mathbf{M}''(x_k))\right|\right) \leq C(\Delta X)^{2k},\end{aligned}$$

where we used eq. (32) in the last step. Moreover, it holds due to orthogonality and eq. (32) that

$$\begin{aligned}|\mathbf{P}_{[2k]}(x_k)\cdot\mathbf{P}'_{[2k]}(x_k)| &= |\mathbf{M}_k\cdot(\mathbf{P}'_{[2k]}(x_k) - \mathbf{M}'(x_k))| \\ &\leq |\mathbf{P}'_{[2k]}(x_k) - \mathbf{M}'(x_k)| \leq C(\Delta X)^{2k}.\end{aligned}$$

It therefore follows that

$$\begin{aligned}&|\mathbf{Q}''_{[2k]}(x_k) - \mathbf{P}''_{[2k]}(x_k)| \\ &\quad\leq 2\left|(\mathbf{P}_{[2k]}(x_k)\cdot\mathbf{P}'_{[2k]}(x_k))\mathbf{P}'_{[2k]}(x_k)\right| + 3\left|\left(\mathbf{P}_{[2k]}(x_k)\cdot\mathbf{P}'_{[2k]}(x_k)\right)^2\mathbf{P}_{[2k]}(x_k)\right| \\ &\qquad + \left|\left(|\mathbf{P}'_{[2k]}(x_k)|^2 + \mathbf{P}''_{[2k]}(x_k)\cdot\mathbf{P}_{[2k]}(x_k)\right)\mathbf{P}_{2k}(x_k)\right| \leq C(\Delta X)^{2k},\end{aligned}$$

which by eq. (32) implies that $\mathbf{Q}''_k(x_k)$ is a $2k$-th order approximation to $\mathbf{M}''$ (but not a standard linear central finite difference approximation). □

We can conclude that when using either $\mathbf{P}_{[2k]}$ or $\mathbf{Q}_{[2k]}$ to obtain initial data $\mathbf{m}_{\text{init}}$ for the HMM micro problem, then $\partial_{xx}\mathbf{m}_{\text{init}}(0)$ is a $2k$-th order approximation to $\partial_{xx}\mathbf{M}(x_k, t_j)$, where x_k is the grid point in the middle of the interpolation stencil. Similarly, in two space dimensions, $\partial_{xy}\mathbf{m}_{\text{init}}(0)$ and $\partial_{yy}\mathbf{m}_{\text{init}}(0)$ are $2k$-th order approximations to $\partial_{xy}\mathbf{M}(x_k, t_j)$ and $\partial_{yy}\mathbf{M}(x_k, t_j)$. Thus, the discretization error, which corresponds to the interpolation error, is

$$E_{\text{disc}} = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_{\text{init}}(0)\mathbf{A}^H) - \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{M}(x_k, t_j)\mathbf{A}^H) = C(\Delta X)^{2k}. \tag{36}$$

Both $\mathbf{P}_{[2k]}$ and $\mathbf{Q}_{[2k]}$ hence satisfy the consistency requirement eq. (30) for the initial data. However, only $\mathbf{Q}_{[2k]}$ is normalized, therefore this is what we choose subsequently. Typically, the difference between the approximations is only rather small, though, as shown in the following numerical example.

Numerical example

As an example, consider the initial data for a micro problem on a 2D domain of size 10ε in each space dimension, obtained by interpolation from the macro initial data of (EX2) and (EX3).

We first investigate the maximal deviation of the length of $\mathbf{m}_{\text{init}}$ from one when using $\mathbf{P}_{[2]}$ and $\mathbf{P}_{[4]}$ to obtain the initial data. In Figure 8, this error is shown for varying ΔX and different values of ε.

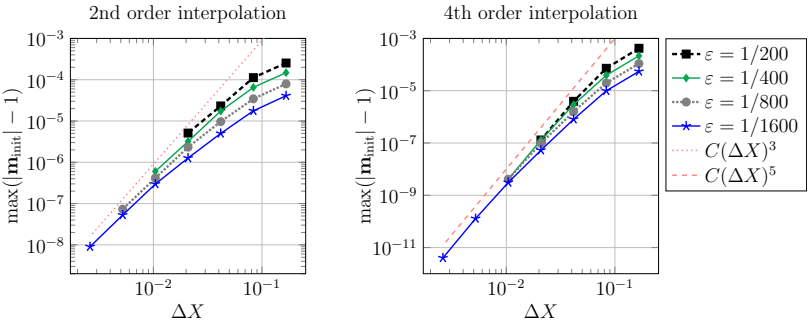


Fig. 8: Maximum norm deviation in polynomial interpolation initial data $\mathbf{P}_{[2]}$ (left) and $\mathbf{P}_{[4]}$ (right) from one given a micro domain size of 10ε, for several values of ε. 2D problem with macro initial data as in (EX3) and (EX2). Only values where $2k\Delta X > 10\varepsilon$ are plotted to avoid extrapolation.

One can observe that the deviation decreases as the macro step size ΔX decreases. Moreover, especially for high ΔX values, smaller ε result in smaller deviations. This is due to the fact that a smaller ε corresponds to a smaller micro domain, around x_k. The maximum possible norm deviation is only attained further away from x_k. In the limit, as $\varepsilon \to 0$, the norm deviation vanishes.

Next, we examine the difference between $\mathbf{P}_{[2k]}$ and $\mathbf{Q}_{[2k]}$ and the order of accuracy of the resulting approximations to $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{M} \mathbf{A}^H)$. In the left subplot in Figure 9, the error E_{disc} is shown for several values of ΔX and for second and fourth order interpolation. The expected convergence rates of $(\Delta X)^{2k}$ can be observed for both approximations, based on $\mathbf{P}_{[2k]}$ and $\mathbf{Q}_{[2k]}$. In the right

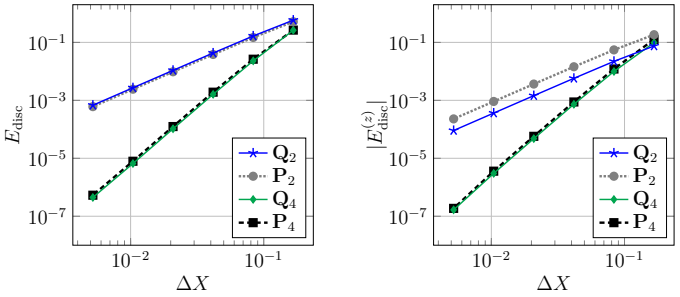


Fig. 9: Interpolation error E_{disc} with and without normalization, i.e. using $\mathbf{Q}_{[2k]}$ and $\mathbf{P}_{[2k]}$, when varying macro grid spacing ΔX. 2nd and 4th order interpolation. Left: norm of the error between approximated and actual effective field for (EX2). Right: z-component only.

subplot, only the difference between the z-components of $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{m}_{\text{init}} \mathbf{A}^H)$ and $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{M} \mathbf{A}^H)$ is considered to emphasize the fact that while $\mathbf{P}_{[2k]}$ and $\mathbf{Q}_{[2k]}$ result in approximations of the same order of accuracy, they do not give the same approximation.

4.2 Boundary conditions

In this section, the issue of boundary conditions for the HMM micro problem is discussed. In the case of a periodic material coefficient, as considered for the estimate in Theorem 2, the micro problem would ideally be solved on the whole domain Ω with periodic boundary conditions, even though the resulting solution is only averaged over a small domain $[-\mu, \mu]^d$. This is not a reasonable choice in practice, since the related computational cost is too high. We therefore have to restrict the size of the computational domain for the micro problem and complete it with boundary conditions. Every choice of boundary conditions introduces some error inside the domain in comparison to the whole domain solution since it is not possible to exactly match both "incoming" and "outgoing" dynamics. In Figure 10, the effect of boundary conditions in comparison to the solution on a much larger domain is illustrated for one example 1D micro problem with the setup (EX1). The solution in the micro domain as well as the errors due to two kinds of boundary conditions are plotted: assuming periodicity of $\mathbf{m}^\varepsilon - \mathbf{m}_{\text{init}}$ (middle) and homogeneous Dirichlet boundary conditions for $\mathbf{m}^\varepsilon - \mathbf{m}_{\text{init}}$ (right). In both cases, one can observe errors propagating into the domain from both boundaries as time increases, even though the

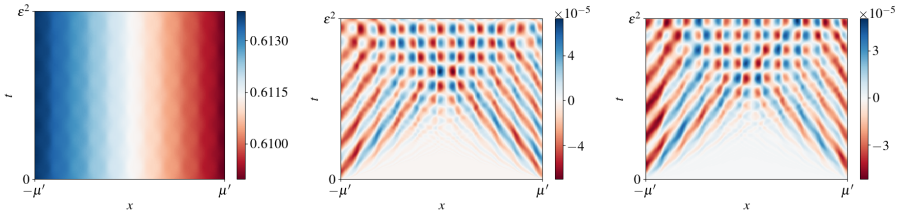


Fig. 10: Example (EX1) with $\alpha = 0.01$, comparison of solution on micro domain $[-\mu', \mu']$, where $\mu' = 5\varepsilon$ and $\varepsilon = 2{\cdot}10^{-3}$, to solution on a 10 times larger domain. Left: x component of the expected solution, middle: error with periodic boundary conditions for $\mathbf{m}^\varepsilon - \mathbf{m}_{\text{init}}$, right: error with Dirichlet boundary condition.

amplitude of the errors is influenced by the type of condition. Since we cannot remove this problem even when considering more involved boundary conditions, we choose to solve the micro problem on a domain $[-\mu', \mu']^d$, for some $\mu' \geq \mu$, with Dirichlet boundary conditions and only average over $[-\mu, \mu]^d$. The size of the domain extension $\mu' - \mu$ together with the time parameter η determine how large the boundary error $E_{\mu'}$ in eq. (29) becomes. For larger values of η, we expect a larger $\mu' - \mu$ to be required to obtain E_μ below a given threshold, since given a longer final time the errors can propagate further into the domain. This is investigated in more detail in the next section.

4.3 Size of micro domain

Here, we investigate how to choose the size of the micro problem domain. There are three important parameters that have to be set, μ and η as in Theorem 2, which determine the size of the averaging domain in space and time, as well as the outer box size μ'. Note that the optimal choice of all three parameters is dependent on the given initial data and the material coefficient.

To determine the influence of the respective parameters, we consider the example (EX2), with a periodic material coefficient, and investigate for one micro problem the error E_{avg} as given in eq. (29). Throughout this section, we consider averaging kernels K, K^0 with $p_x = p_t = 3$ and $q_x = q_t = 7$, based on the experiments in [8]. Typically, $\varepsilon = 1/400$ is used, which results in a value of E_ε that is relatively low compared to other contributions to E_{avg}. Moreover, we assume that E_{num} is small compared to the other error terms and can be neglected. In practice, we here use at least 40 discretization points per ε to achieve this.

Averaging domain size, μ

To begin with, we choose a large value for the computational domain μ', so that

$$E_{\text{avg}} \approx E_\varepsilon + E_\eta + E_\mu.$$

We then vary the averaging parameter μ, which affects the error contribution E_μ, that satisfies eq. (13), repeated here for convenience,

$$E_\mu \leq C \left(\mu^{p_x+1} + \left(\frac{\varepsilon}{\mu} \right)^{q_x+2} \right). \tag{37}$$

Based on the considerations in [8] and Theorem 2, we expect that μ should be chosen to be a multiple of ε, which is the scale of the fast spatial oscillations in the problem. With the given averaging kernel, the first term on the right-hand side in eq. (37) then is small in comparison to the other error contributions and the second term dominates E_μ.

In Figure 11, the development of E_{avg} when increasing μ for (EX2) is shown for several values of η and α. One can observe that as μ is increased, the error decreases rapidly from high initial levels. This is due to the contribution $C \left(\frac{\varepsilon}{\mu} \right)^{q_x+2}$ to E_μ. Once μ becomes sufficiently large, in this example around $\mu \approx 3.5\varepsilon$, the error does not change significantly anymore but stays at a constant level, depending on η and α. Here E_μ no longer dominates the error, which will instead be determined by E_η and E_ε. One can observe that longer times η result in lower overall errors. Furthermore, the errors in the high damping case, $\alpha = 1$, are considerably lower than with $\alpha = 0.1$ or $\alpha = 0.01$. However, note that the required value of μ until the errors no longer decrease is independent of both α and η. In the subsequent investigations, we therefore choose a fixed value of μ which is slightly above this number, thus making sure that E_μ does not significantly influence the overall error observed there.

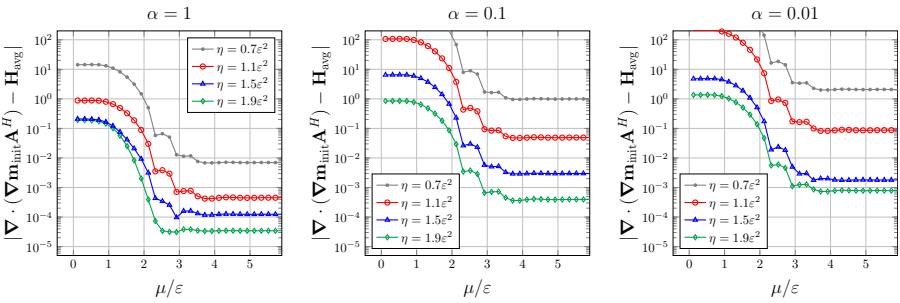


Fig. 11: (EX2): Influence of spatial averaging size μ on the overall error in one micro problem for several values of η and α when $\varepsilon = 1/400$. Kernel parameters $p_x = p_t = 3$ and $q_x = q_t = 7$. The outer box size μ' is chosen sufficiently large to not significantly influence the results.

Full domain size, μ'

Next, we study the effect of the size of the computational domain, which is determined by μ', on the error E_{avg}. We here choose μ large enough so that

$$E_{\text{avg}} \approx E_\varepsilon + E_\eta + E_{\mu'}.$$

We consider the same choices of α and η-values as in the previous example, and vary μ' to investigate $E_{\mu'}$. Note that in contrast to the other error terms, we do not have a model for $E_{\mu'}$. As shown in Figure 12, a value of μ' that is only slightly larger than μ gives a high error, which decreases as μ' is increased, until the same error levels as in Figure 11, determined by E_η and E_ε, are reached. The longer the time interval η considered, the larger the domain has

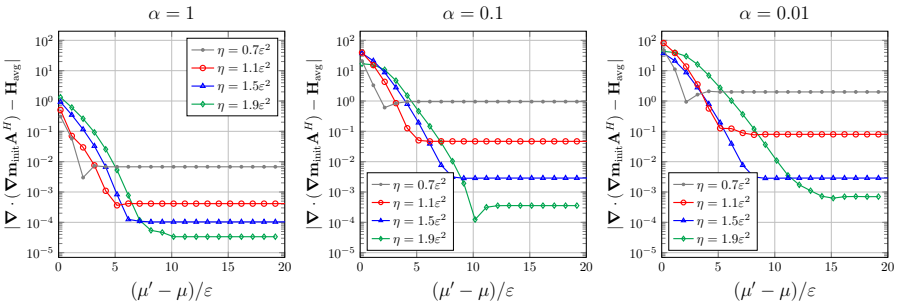


Fig. 12: (EX2): Influence of the extension of the computational domain, $[-\mu', \mu']^d$ beyond the averaging domain, $[-\mu, \mu]^d$, on the overall averaging error E_{avg} in one micro problem for several values of η and α. Here $\mu = 3.9\varepsilon$ and $\varepsilon = 1/400$.

to be chosen to reduce the boundary error $E_{\mu'}$ such that it no longer dominates E_{avg}. This is due to the fact that the boundary error propagates further into the domain the longer time passes. We can furthermore observe that larger α results in somewhat faster convergence of the error for higher values of η.

Length of time interval η

Finally, we consider the influence of η and the corresponding error contribution E_η to the averaging error E_{avg} as given in eq. (29). Based on Theorem 2, we have

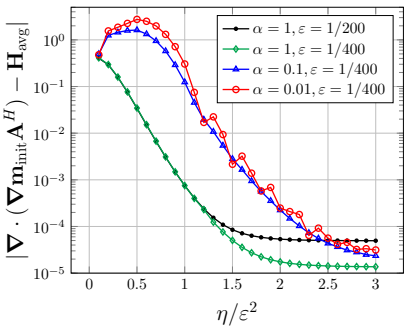
$$E_\eta \leq C_\mu \left(\eta^{p_t+1} + \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1} \right), \tag{38}$$

repeated here for convenience. We consider $\eta \sim \varepsilon^2$. With the given choice of averaging kernel, with $p_t = 3$, the first term in eq. (38) is small compared to

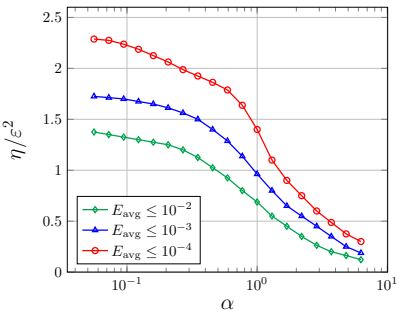
the second one. We choose the parameters μ and μ' such that

$$E_{\text{avg}} \approx E_\varepsilon + E_\eta$$

and vary η. In Figure 13a, one can then observe that higher values of η result in lower errors, since the second term on the right hand side in eq. (38) decreases as η increases. This matches with the error behavior depicted in Figure 11 and Figure 12. Figure 13a furthermore shows that the error eventually saturates at a certain level, corresponding to E_ε. Comparing the errors for $\alpha = 1$ with $\varepsilon = 1/200$ and $\varepsilon = 1/400$, one finds that the respective E_ε differ by a factor of approximately four, which indicates that $E_\varepsilon \leq C\varepsilon^2$ here, as was previously observed in [8]. The different cases of α considered in Figure 13a have a similar overall behavior of the error, but for high damping, $\alpha = 1$, the development happens for considerably lower values of η than in the other cases. Moreover, in case of $\alpha = 0.01$, we observe some oscillations as the error decreases.



(a) Averaging error E_{avg} when varying η.



(b) Time required until E_{avg} falls below given thresholds when varying α.

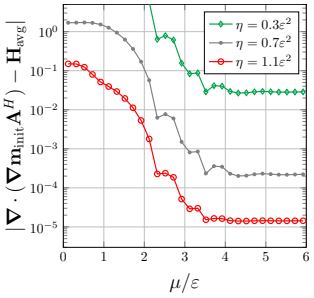
Fig. 13: (EX2): Influence of the time averaging length η on the overall error in one micro problem. Here $\mu = 3.9\varepsilon$ and μ' is chosen sufficiently big to not significantly change the results. Kernel parameters $p_t = 3$ and $q_t = 7$.

In Figure 13b, we further investigate the influence of the damping parameter on the time η it takes for the averaging error E_{avg} to fall below certain given thresholds. One can clearly observe that high damping reduces the required time to reach all three considered error levels. This indicates that the introduction of artificial damping in the micro problem can help to significantly reduce computational cost, since a shorter final time also implies a smaller computational domain as explained in the previous section. However, since $\alpha \gg 1$ results in a seriously increased number of time steps necessary to get a stable solution, as discussed in Section 3.2, we conclude that choosing α around one is most favorable. Note that as discussed in Section 2.3 in connection with

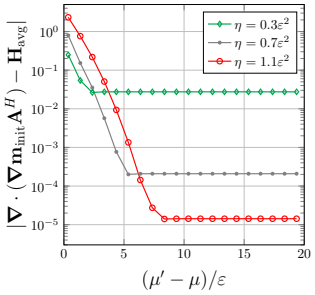
Theorem 2, it is proved for periodic material coefficients that choosing a higher value of α does not change the quantity $\mathbf{H}_{\text{avg}}$ which is approximated in the upscaling process.

Example (EX3)

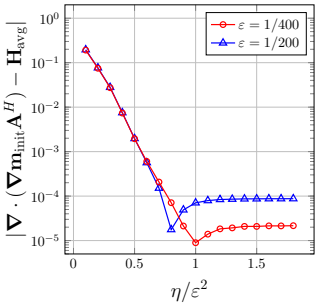
To support the considerations regarding the choice of micro parameters, we furthermore study the Landau-Lifshitz problem eq. (2) with the setup (EX3). In (EX3) the material coefficient has a higher average and higher maximum value than in (EX2). This results in a higher "speed" of the dynamics. In Figure 14, the influence of μ, μ', η and α, respectively, on the averaging error E_{avg} are shown for this example.



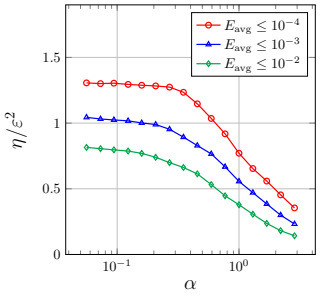
(a) Influence of averaging domain parameter μ, compare to Figure 11.



(b) Influence of computational domain extension $\mu' - \mu$, see Figure 12.



(c) Influence of averaging time η, compare to Figure 13a.



(d) Influence of damping α, compare to Figure 13b.

Fig. 14: (EX3): Influence of micro domain parameters on error E_{avg}. Parameters not explicitly given are chosen to not influence E_{avg} significantly. Moreover, $\varepsilon = 1/400$, and $\alpha = 1$ in (a) - (c).

Qualitatively, the results for (EX3) are the same as (EX2), but some details differ. A notable difference between the examples (EX2) and (EX3) is that the time η required for saturation of the errors is considerably shorter in (EX3), as can be observed when comparing Figure 14c to Figure 13a. An explanation for this is that due to the faster dynamics in (EX3), comparable effects are achieved at shorter times. The ratio between the times η it takes in (EX2) and (EX3), respectively, to reach the level where the error no longer changes matches approximately with the ratio of the maxima of the material coefficients.

Moreover, a slightly larger μ is required to reach the level where the errors saturate in (EX3). The (EX3) saturation errors for a specific value of η are lower, though, due to the fact that the error decreases faster with η as discussed previously.

When it comes to the full size of the domain required for the boundary error to not influence the overall error in a significant way, one can observe somewhat larger required value of μ' in (EX3) when comparing Figure 14b to Figure 12. This is partly due to the fact that the overall error for a given η is lower in (EX3). Moreover, due to the faster dynamics, the errors at time η have propagated further into the domain. As a result, the computational domain has to be chosen approximately the same size in (EX2) and (EX3) to obtain a certain error level, even though the required η is smaller in (EX3).

4.4 Computational Cost

The computational cost per micro problem is a major factor for an efficient HMM implementation. Given a spatial discretization with a certain number of grid points, K, per wave length ε, that is a micro grid spacing $\delta x = \varepsilon/K$, we have in total $N^d = (2K\mu'/\varepsilon)^d$ micro grid points. The time step size for the micro time integration has to be chosen as $\delta t \leq C_{\text{stab},\alpha}\delta x^2$, hence the number of time steps becomes $M \geq C^{-1}_{\text{stab},\alpha}K^2\eta/\varepsilon^2$. It then holds for the computational cost per micro problem that

$$\text{micro cost} \sim MN^d \sim \frac{1}{C_{stab,\alpha_{\text{micro}}}}\frac{\eta}{\varepsilon^2}\left(\frac{\mu'}{\varepsilon}\right)^d K^{2+d}. \tag{39}$$

It is important to note that due to the choice of parameters $\mu' \sim \varepsilon$ and $\eta \sim \varepsilon^2$, the computational cost per micro problem is independent of ε. This makes it possible to use HMM also for problems where the computational cost of other approaches, resolving the fast oscillations, becomes tremendously high.

In general, choosing higher values for η, μ' or K results in higher computational cost per micro problem. We therefore aim to choose these values as low as possible without negatively affecting the overall error. For the numerical examples shown in Section 5, we find that selecting $K \approx 10$ gives a reasonable compromise between accuracy and computational cost. The overall cost is determined by the cost per micro problem and the choice of the macro

discretization size ΔX,

$$\text{cost} \sim \frac{1}{C_{\text{stab},\alpha_{\text{macro}}}} (\Delta X)^{-(2+d)} \text{micro cost}.$$

This shows the importance of choosing ΔX relatively large, wherefore it is advantageous that the overall method proposed based on MPEA is fourth order accurate, as discussed in Section 3.3. Since all the HMM micro problems are independent of each other, it is moreover very simple to parallelize their computations. This can be an effective way to reduce the overall run time of the method.

4.5 Choice of overall setup

We now aim to give an overview of the overall choice of HMM setup, based on the discussions in Section 3 and Sections 4.1 to 4.3. Furthermore, we study the resulting errors.

To begin with, we discuss the overall setup for the micro problem. According to eq. (28), the approximation error in $\mathbf{H}_{\text{avg}}$ is

$$E_{\text{approx}} = E_{\text{avg}} + E_{\text{disc}},$$

where E_{disc} is determined by the macro discretization size ΔX as given in eq. (36). For a given ΔX, we therefore aim to choose the parameters μ, μ' and η so that E_{avg} matches E_{disc}. Further reducing E_{avg} only increases the computational cost per micro problem without significantly improving the overall error.

In Table 1, we hence suggest choices for η and μ' in the example setups (EX2) and (EX3) with $\Delta X = 1/(12 \cdot 2^i)$, $i = 0, 1, 2$, such that E_{avg} is slightly below the corresponding values for E_{disc} when using fourth order interpolation. The averaging parameter μ is fixed to a value such that E_μ does not significantly increase E_{avg} but not much higher. This helps to reduce the number of parameters to vary. Moreover, choosing lower μ results in a rather steep increase of E_{avg} in comparison to how much the computational cost is reduced.

Note that the specific values of the discretization error depend on the given macro solution and macro location. In this example, we consider the macro initial data and the micro problem solved to obtain $\mathbf{H}_{\text{avg}}$ at macro location $(0, 0)$. We furthermore choose $\alpha = 1$ here to make it simple to compare the suggestions to the values shown in Figures 11, 12 and 13a as well as Figure 14. The optimal choice for α would be slightly higher.

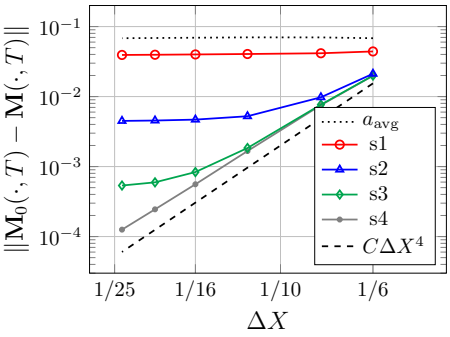
Based on the values in Table 1, one can conclude that the required sizes of the computational domains are very similar between the considered examples. When scaled by the maximum of the respective material coefficients, also the suggested final times are comparable.

Next, we further test the influence of the micro domain setup and corresponding E_{avg} on the overall error. For this purpose, we consider (EX2) on a

		(EX2)			(EX3)		
ΔX	E_{disc}	η/ε^2	$a_{\max}\eta/\varepsilon^2$	$(\mu'-\mu)/\varepsilon$	η/ε^2	$a_{\max}\eta/\varepsilon^2$	$(\mu'-\mu)/\varepsilon$
1/12	$2.3\cdot10^{-2}$	0.7	1.09	4	0.4	1.02	3
1/24	$1.6\cdot10^{-3}$	1	1.56	6	0.6	1.53	5
1/48	$1\cdot10^{-4}$	1.4	2.18	8	0.8	2.05	7

Table 1: Example micro problem setups with $\alpha = 1$, where $\mu = 3.9\varepsilon$ for (EX2) and $\mu = 4.2\varepsilon$ for (EX3) .

unit square domain with periodic boundary condition, $\alpha = 0.01$ in the original problem and a final time $T = 0.1$. We use artificial damping and set $\alpha = 1.2$ in the micro problem. As in the previous examples, we choose averaging kernel parameters $p_x = p_t = 3$ and $q_x = q_t = 7$ and let $\varepsilon = 1/400$. We again fix $\mu = 3.9\varepsilon$, and use HeunP as the micro solver as proposed in Section 3. . We then run HMM with MPEA for the macro time stepping to approximate the homogenized reference solution $\mathbf{M}_0$ at time T for varying ΔX. Four different



	η/ε^2	μ'	E_{avg}
s1	0.15	4	$\approx 3\cdot10^{-1}$
s2	0.45	5.5	$\approx 3\cdot10^{-2}$
s3	0.7	7.5	$\approx 3\cdot10^{-3}$
s4	1	10	$\approx 3\cdot10^{-4}$

Micro problem setups for (EX2) (with $\alpha = 1.2$ and $\mu = 3.9$) and resulting E_{avg} in micro problem at macro location $x_0 = (0, 0)$.

Fig. 15: L^2-norm of difference between HMM solution $\mathbf{M}$ and homogenized reference solution $\mathbf{M}_0$ at time $T = 0.1$ for (EX2).

combinations of η and μ' are used for the micro problem, referred to as (s1)-(s4). The resulting L^2-norms of the errors $\mathbf{M}_0 - \mathbf{M}$ are shown in Figure 15, together with the error one obtains when using the average a_{avg} of the material coefficient a^ε to approximate $\mathbf{A}^H$ when solving eq. (3). Using a_{avg} can be seen as a naive approach to dealing with the fast oscillations in the material coefficient. It does in general not result in good approximations. The corresponding error is included here to give a baseline for the relevance of the HMM solutions.

We find that with the micro problem setup (s1), corresponding to a rather high averaging error, HMM results in a solution that is only slightly better than the one that is obtained using the average of the material coefficient. However, when applying setups with lower averaging errors, lower overall errors are achieved. In particular, with (s4), the setup with the lowest considered averaging error, the overall error in Figure 15 is determined by the error E_{disc}, proportional to $(\Delta X)^4$ since fourth order interpolation is used to obtain the

initial data for the micro problems. For the other two setups, (s2) and (s3), the overall errors saturate at levels somewhat lower than the respective values of E_{avg}, corresponding to e_{HMM} in eq. (27). Note that this saturation occurs for relatively high values of ΔX.

For matters of completeness, in Figure 16, also the H^1-norm of the error in the HMM solution compared to the homogenized solution $\mathbf{M}_0$ is shown for the same setups as in Figure 15. The observed errors behave very similar to the L^2-case. Note, though, that the H^1-norm of the homogenization error, $\|\mathbf{M}^\varepsilon - \mathbf{M}_0\|_{H^1}$, is only known to be bounded [7], it is in general not small, in contrast to $\|\mathbf{M}^\varepsilon - \mathbf{M}_0\|_{L^2}$ which is $\mathcal{O}(\varepsilon)$.

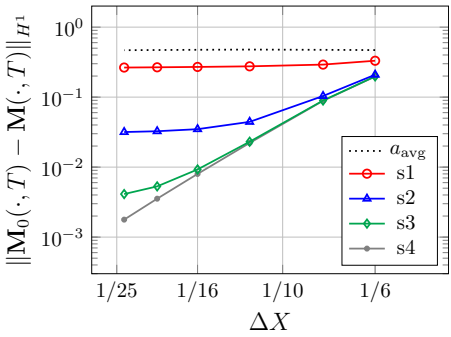


Fig. 16: H^1-norm of difference between HMM solution $\mathbf{M}$ and homogenized reference solution $\mathbf{M}_0$ at time $T = 0.1$ for (EX2).

5 Further numerical examples

To conclude this article, we consider several numerical examples with material coefficients that are not fully periodic. Those cases are not covered by the theorems in Section 2, however, the HMM approach still results in good approximations. In the 2D examples, we again include the solution obtained when using a (local) average of the material coefficient as an approximation to the effective coefficient to stress the relevance of the HMM solutions. As for the periodic examples, we use artificial damping in the micro problem.

Locally periodic 1D example

We first consider a one-dimensional example with material coefficient

$$a^\varepsilon(x) = 1.1 + \tfrac{1}{4}\sin(2\pi x + 1.1) + \tfrac{1}{2}\sin(2\pi x/\varepsilon). \tag{40}$$

This coefficient is locally periodic. We consider eq. (2) with this coefficient and $\varepsilon = 1/400$, $\alpha = 0.01$ on the unit interval with periodic boundary conditions. A comparison between the solution $\mathbf{M}^\varepsilon$ at time $T = 0.1$, obtained using a

direct numerical simulation resolving the ε-scale, and corresponding HMM approximation on a coarse grid with $\Delta X = 1/25$ is shown in Figure 17. Here the HMM parameters are chosen to be $\mu = 3.9\varepsilon$, $\mu' = 8\varepsilon$ and $\eta = 0.9\varepsilon^2$. Artificial damping with $\alpha = 1.2$ is used for the micro problem. The averaging kernel parameters are again $p_x = p_t = 3$ and $q_x = q_t = 7$. For the direct

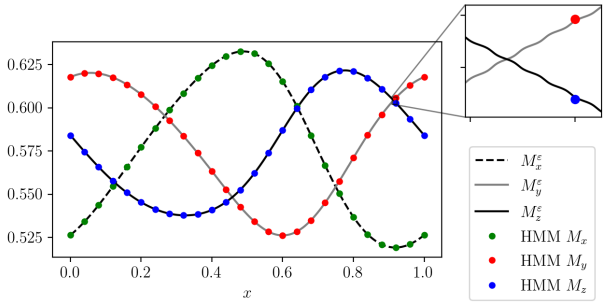


Fig. 17: HMM solution to eq. (2) with a^ε as in eq. (40) with $\varepsilon = 1/400$ and corresponding $\mathbf{M}^\varepsilon$ obtained using direct simulation resolving ε, at $T = 0.1$.

simulation, we use $\Delta x = 1/6000$, which corresponds to 15 grid points per ε, and MPEA for time integration.

One can clearly observe that the HMM solution is very close to the solution obtained with a direct simulation resolving ε. The corresponding discrete L^2 norm of the error is approximately $\|\mathbf{M}_0 - \mathbf{M}_{\text{HMM}}\|_{L^2} \approx 10^{-4}$. Moreover, note that in this example the computation time for HMM is about 20 seconds[1], while the direct simulation takes almost two hours.

Quasi-periodic 2D example

Next, eq. (2) is solved in two space dimensions and with material coefficient

$$a^\varepsilon(x) = (1 + 0.25\sin(2\pi x_1/\varepsilon))(1 + 0.25\sin(2\pi x_2/\varepsilon) + 0.25\sin(2\pi r x_2/\varepsilon)), \tag{41}$$

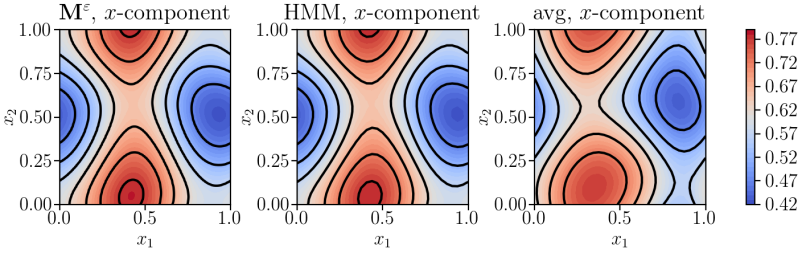
where $r = 1.41$ as an approximation to $\sqrt{2}$. This coefficient is periodic in x_1-direction but not in x_2-direction. If we choose $\varepsilon = 0.01$, though, it is periodic also in x_2 direction over the whole domain $[0,1]^2$ but not on the micro domains. The initial data is set as in (EX2).

To make direct numerical simulation feasible, we consider the case $\varepsilon = 0.01$ and set $\Delta x = 1/1500$ in the direct simulation. For HMM, the micro problem parameters are chosen to be $\mu = 6.5\varepsilon$, $\eta = 0.7\varepsilon^2$ and $\mu' = 9.25\varepsilon$. We use again averaging kernels with $p_x = p_t = 3$ and $q_x = q_t = 7$ as well as artificial

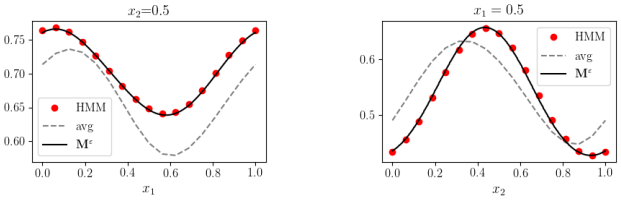
[1] on a computer with Intel i7-4770 CPU at 3.4 GHz

damping with $\alpha = 1.2$ in the micro problem. On the macro scale, $\Delta X = 1/16$ and $\alpha = 0.01$. The final time is set to $T = 0.2$.

In Figure 18a, the x-components of $\mathbf{M}^\varepsilon$, obtained using a direct simulation resolving ε and a HMM solution to eq. (2) with material coefficient eq. (41) are shown. Moreover, the solution obtained when simply using the average of $a^\varepsilon(x)$ as an approximation is included. One can observe that the HMM solution captures the characteristics of the overall solution well, while the approach with an averaged coefficient does not. To further stress this, cross sections of the respective solutions at $x_1 = 0.5$ and $x_2 = 0.5$ are shown in Figure 18b.



(a) Contours



(b) Cross sections

Fig. 18: Quasi-periodic example, with $\varepsilon = 0.01$, $T = 0.2$ and $\alpha = 0.01$.

Despite the choice of a rather high ε-value, $\varepsilon = 0.01$, the direct simulation of this problem took about 5 days. In comparison, the computational time of HMM was about 4 hours[2], which is independent of ε.

Locally periodic 2D example

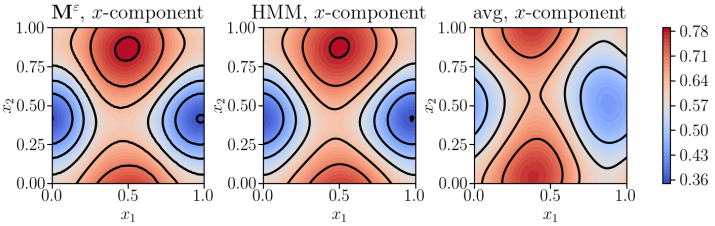
Finally, we consider a locally periodic 2D example with material coefficient

$$a^\varepsilon(x) = 0.25 \exp\left(-\cos(2\pi(x_1 + x_2)/\varepsilon) + \sin(2\pi x_1/\varepsilon)\cos(2\pi x_2)\right). \tag{42}$$

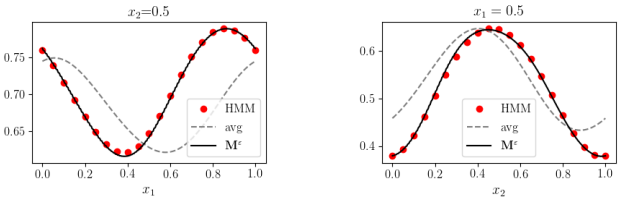
In this example, we set $\alpha = 0.1$ and choose a final time $T = 0.05$.

[2]on a computer with Intel Xeon E5-2637 v3 CPU at 3.50GHz

The HMM parameters are set to $\mu = 5\varepsilon$, $\mu' = 9.75\varepsilon$ and $\eta = 1.6\varepsilon^2$. Note that the average and maximum of this coefficient is lower than for the previously considered ones, hence a longer time η is chosen. Again $\alpha = 1.2$ in the micro problem. Initial data and averaging parameters are set as in the previous example. Direct simulation solution $\mathbf{M}^\varepsilon$, HMM approximation and a solution based on local averages of a^ε are shown in Figure 19. Also for this problem HMM captures the characteristics of the solution well. The discrete L^2 norm of the error with the chosen micro setup is approximately $5 \cdot 10^{-3}$. In contrast, the averaging based solution does not capture the correct characteristics.



(a) Contours



(b) Cross sections

Fig. 19: Locally periodic example, with $\varepsilon = 0.01$, $T = 0.05$ and $\alpha = 0.1$.

6 Conclusion and future work

In this paper, we presented a heterogeneous multiscale method for the Landau-Lifshitz equation with highly oscillatory material coefficient in a periodic setting. A simplified setup, only taking into account the so-called exchange interaction was considered. We showed convergence of relevant errors for periodic material coefficients and obtain solutions which capture the correct characteristics for more general cases. To get a more physically realistic setting, the model should be extended to also include anisotropy, external field and demagnetization. Furthermore, additional boundary conditions should be studied. In particular, in physics and materials science it is common to consider homogeneous Neumann boundary conditions. A step in this direction

is taken in [32], where a finite element method for a similar Landau-Lifshitz problem in a somewhat more realistic setting is discussed.

Furthermore, in [33], a basis-representation technique which drastically reduces the number of micro problems that have to be solved in HMM was introduced for linear PDEs. It would be very interesting to investigate whether those ideas can be extended to the non-linear problem considered here. This would be very beneficial for the HMM scheme discussed in this paper.

Acknowledgments

The authors would like to thank Professor Gunilla Kreiss for useful discussions in the process of this work.

Declarations

The research leading to these results received funding from the Swedish Research Council under grant no 2017-04579 which is gratefully acknowledged. The authors have no competing interests to declare that are relevant to the content of this article.

Data availablility statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

[1] Hamdache, K.: Homogenization of layered ferromagnetic media. Ecole Polytechnique, Centre de Mathématiques Appliquées, R. I. No 495, `http://www.cmap.polytechnique.fr/ preprint/repository/495.ps` (gzip archive) (2002)

[2] Alouges, F., De Bouard, A., Merlet, B., Nicolas, L.: Stochastic homogenization of the Landau-Lifshitz-Gilbert equation. Stochastics and Partial Differential Equations: Analysis and Computations, 1–30 (2021)

[3] Santugini-Repiquet, K.: Homogenization of ferromagnetic multilayers in the presence of surface energies. ESAIM: Control, Optimisation and Calculus of Variations **13**(2), 305–330 (2007)

[4] Choquet, C., Moumni, M., Tilioua, M.: Homogenization of the Landau-Lifshitz-Gilbert equation in a contrasted composite medium. Discrete & Continuous Dynamical Systems-S **11**(1), 35 (2018)

[5] Arjmand, D., Engblom, S., Kreiss, G.: Temporal upscaling in micromagnetism via heterogeneous multiscale methods. J. Comput. Appl. Math. **345**, 99–113 (2019)

[6] Arjmand, D., Kreiss, G., Poluektov, M.: Atomistic-continuum multiscale modeling of magnetization dynamics at non-zero temperature. Adv. Comput. Math **44**, 1119–1151 (2018)

[7] Leitenmaier, L., Runborg, O.: On homogenization of the Landau–Lifshitz equation with rapidly oscillating material coefficient. Communications in Mathematical Sciences **20**(3), 653–694 (2022)

[8] Leitenmaier, L., Runborg, O.: Upscaling errors in heterogeneous multiscale models for the Landau-Lifshitz equation. preprint (2021)

[9] E, W., Engquist, B., *et al.*: The Heterognous Multiscale Methods. Communications in Mathematical Sciences **1**(1), 87–132 (2003)

[10] Abdulle, A., E, W., Engquist, B., Vanden-Eijnden, E.: The Heterogeneous Multiscale Method. Acta Numerica **21**, 1–87 (2012)

[11] E, W., Engquist, B., Li, X., Ren, W., Vanden-Eijnden, E.: Heterogeneous Multiscale Methods: a review. Communications in computational physics **2**(3), 367–450 (2007)

[12] Engquist, B., Tsai, Y.-H.: Heterogeneous multiscale methods for stiff ordinary differential equations. Mathematics of computation **74**(252), 1707–1742 (2005)

[13] Arjmand, D., Runborg, O.: Analysis of heterogeneous multiscale methods for long time wave propagation problems. Multiscale Modeling & Simulation **12**(3), 1135–1166 (2014)

[14] Arjmand, D., Runborg, O.: A time dependent approach for removing the cell boundary error in elliptic homogenization problems. Journal of Computational Physics **314**, 206–227 (2016)

[15] García-Cervera, C.J.: Numerical micromagnetics: A review. Bol. Soc. Esp. Mat. Apl. 39, 103–135 (2007)

[16] Cimrák, I.: A survey on the numerics and computations for the Landau-Lifshitz equation of micromagnetism. Archives of Computational Methods in Engineering **15**(3), 1–37 (2007)

[17] Baňas, L.: Numerical methods for the Landau-Lifshitz-Gilbert equation. In: International Conference on Numerical Analysis and Its Applications, pp. 158–165 (2004). Springer

[18] Wang, X.-P., Garcıa-Cervera, C.J., E, W.: A Gauss–Seidel projection method for micromagnetics simulations. Journal of Computational Physics **171**(1), 357–372 (2001)

[19] Li, P., Xie, C., Du, R., Chen, J., Wang, X.-P.: Two improved Gauss-Seidel projection methods for Landau-Lifshitz-Gilbert equation. Journal of Computational Physics **401**, 109046 (2020)

[20] E, W., Wang, X.-P.: Numerical methods for the Landau-Lifshitz equation. SIAM journal on numerical analysis, 1647–1665 (2001)

[21] d'Aquino, M., Serpico, C., Miano, G.: Geometrical integration of Landau–Lifshitz–Gilbert equation based on the mid-point rule. Journal of Computational Physics **209**(2), 730–753 (2005)

[22] d'Aquino, M., Serpico, C., Miano, G., Mayergoyz, I., Bertotti, G.: Numerical integration of Landau–Lifshitz–Gilbert equation based on the midpoint rule. Journal of applied physics **97**(10), 10–319 (2005)

[23] Mentink, J., Tretyakov, M., Fasolino, A., Katsnelson, M., Rasing, T.: Stable and fast semi-implicit integration of the stochastic Landau–Lifshitz equation. Journal of Physics: Condensed Matter **22**(17), 176001 (2010)

[24] Serpico, C., Mayergoyz, I., Bertotti, G.: Numerical technique for integration of the Landau–Lifshitz equation. Journal of Applied Physics **89**(11), 6991–6993 (2001)

[25] Lewis, D., Nigam, N.: Geometric integration on spheres and some interesting applications. Journal of Computational and Applied Mathematics **151**(1), 141–170 (2003)

[26] Krishnaprasad, P.S., Tan, X.: Cayley transforms in micromagnetics. Physica B: Condensed Matter **306**(1-4), 195–199 (2001)

[27] Leitenmaier, L.: Analysis and numerical methods for multiscale problems in magnetization dynamics. PhD thesis, KTH. (2021)

[28] Wei, D.: Micromagnetics and Recording Materials, 1st edn. Springer, Berlin, Heidelberg (2012)

[29] Mayergoyz, I.D., Bertotti, G., Serpico, C.: Nonlinear Magnetization Dynamics in Nanosystems. Elsevier, Oxford, UK (2009)

[30] García-Cervera, C.J., *et al.*: Improved Gauss-Seidel projection method for micromagnetics simulations. IEEE transactions on magnetics **39**(3), 1766–1770 (2003)

[31] Peiró, J., Sherwin, S.: Finite Difference, Finite Element and Finite Volume Methods for Partial Differential Equations. Handbook of Materials Modeling: Methods, pp. 2415–2446. Springer, Dordrecht (2005)

[32] Leitenmaier, L., Nazarov, M.: A finite element based Heterogeneous

Multiscale Method for the Landau-Lifshitz equation. arXiv preprint arXiv:2111.11197 (2021)

[33] Engquist, B., Holst, H., Runborg, O.: Multi-scale methods for wave propagation in heterogeneous media. Commun. Math. Sci. **9**, 33–56 (2011)