UPSCALING ERRORS IN HETEROGENEOUS MULTISCALE METHODS FOR THE LANDAU-LIFSHITZ EQUATION

LENA LEITENMAIER[*] AND OLOF RUNBORG[†]

Abstract. In this paper, we consider several possible ways to set up Heterogeneous Multiscale Methods for the Landau-Lifshitz equation with a highly oscillatory diffusion coefficient, which can be seen as a means to modeling rapidly varying ferromagnetic materials. We then prove estimates for the errors introduced when approximating the relevant quantity in each of the models given a periodic problem, using averaging in time and space of the solution to a corresponding micro problem. In our setup, the Landau-Lifshitz equation with highly oscillatory coefficient is chosen as the micro problem for all models. We then show that the averaging errors only depend on ε, the size of the microscopic oscillations, as well as the size of the averaging domain in time and space and the choice of averaging kernels.

Key words. Heterogeneous Multiscale Methods; Micromagnetics; Magnetization Dynamics;

AMS subject classifications. 65M15; 35B27; 78M40

1. Introduction. In micromagnetics, the evolution of the magnetization within a ferromagnet is described by the Landau-Lifshitz (LL) equation [16]. In this paper, we consider a simplification of the deterministic version of this equation, where we only take into account the exchange interaction between magnetic moments and neglect other contributions influencing the magnetization, such as anisotropy, temperature and external field. We consider a ferromagnet with a rapidly varying material, which we model by introducing a material coefficient $a^\varepsilon(x)$, where $\varepsilon \ll 1$ represents the spatial scale of the finest variations.. One example could be a composite, consisting of two different materials with different interaction behavior and layers of thickness ε. According to this simplified model, the partial differential equation determining the evolution of the magnetization $\mathbf{M}^\varepsilon(x,t)$ then is

$$
\begin{aligned}
(1.1) \quad & \partial_t \mathbf{M}^\varepsilon(x,t) = -\mathbf{M}^\varepsilon(x,t) \times \mathcal{L}\mathbf{M}^\varepsilon(x,t) - \alpha \mathbf{M}^\varepsilon(x,t) \times \mathbf{M}^\varepsilon(x,t) \times \mathcal{L}\mathbf{M}^\varepsilon(x,t), \\
& \mathbf{M}^\varepsilon(x,0) = \mathbf{M}_{\text{init}}(x),
\end{aligned}
$$

where $\mathbf{M}_{\text{init}}(x)$ is a smooth function with values in $\mathbb{R}^3$ such that $|\mathbf{M}_{\text{init}}(x)| = 1$, and $0 < \alpha \le 1$ a damping coefficient. In this model, the effective field is given by

$$
(1.2) \qquad \mathcal{L}\mathbf{M}^\varepsilon := \boldsymbol{\nabla} \cdot (a^\varepsilon(x) \boldsymbol{\nabla} \mathbf{M}^\varepsilon(x,t)).
$$

Here the coefficient a^ε influences the overall behavior of the magnetization significantly. A very similar model was first introduced in [13] and used recently in [2]. Also in for example [20] and [8], related approaches are applied.

When solving (1.1) numerically, one would have to resolve the ε-scale in order to get a correct result. However, the resulting amount of computational work is infeasible for small ε. Instead of solving (1.1), one therefore in many cases considers solutions to a corresponding effective equation instead, which capture the correct magnetization behavior on a coarse scale but do not resolve the ε-scale. For periodic problems, one can apply techniques from classic homogenization theory, [9], [7], to obtain such a

---

*Department of Mathematics, KTH, Royal Institute of Technology, Stockholm, Sweden (lenalei@kth.se).

†Department of Mathematics, KTH, Royal Institute of Technology, Stockholm, Sweden (olofr@kth.se).

homogenized solution $\mathbf{M}_0$ as well as correction terms as shown in [17]. When aiming to deal with somewhat more general coefficients, though, it can be advantageous to instead use numerical methods to approximate the homogenized solution. This can be done using multiscale methods like equation free methods [14] or heterogeneous multiscale methods (HMM) [21], [22], [1]. The basic idea of HMM is to combine a coarse scale macro model, that involves some unknown quantity, with micro problems that are solved on a short time interval and small domain only. By choosing the size of these proportional to ε the computational cost becomes virtually independent of ε. In the so-called upscaling process, the solution from the micro problem is then averaged to obtain the quantity that is needed to complete the macro model. This is the approach that we consider in this paper. In particular, we choose three different HMM macro models for (1.1). For the case of a periodic material coefficient, we then investigate the upscaling error for each of the models, in order to get an understanding of what are good ways to set up HMM for this problem. We come to the conclusion that all three models give very similar results and can thus be valid choices for HMM setups. Which model to choose can thus mostly be based on advantages related to the numerical implementation.

HMM has previously been applied to a Landau-Lifshitz problem in [3], [4]. However, in these articles, the authors do not consider the case with a material coefficient that is highly oscillatory in space, (1.1), but instead a highly oscillatory external field with temporal oscillations.

In the remainder of this section, we shortly introduce some of the notation that is used in the following. We furthermore describe the homogenized solution for (1.1) with a periodic material coefficient as derived in [17], which will subsequently act as a reference. We continue in Section 2 by describing the concept of heterogeneous multiscale methods as well as the models considered. In Section 3, estimates for the homogenized solution and the corresponding correctors to the HMM micro problem are stated to provide the basis that is required for the subsequent derivations. We moreover add an explicit description of a particular correction term. In Section 4, we derive several lemmas regarding the averaging required for numerical homogenization. These lay the ground for the error estimates for the different upscaling-models, which are given in Section 5 and constitute the main result of this paper. Finally, in Section 6 we present several numerical examples in one and two space dimensions which show the validity of the theoretical estimates.

1.1. Preliminaries. We consider ℓ-periodic solutions to (1.1) in the d-dimensional hypercube $\Omega = [0, \ell]^d$ and time interval $[0, T]$, for some $\ell > 0$. We use Y to denote the d-dimensional unit cell $[0, 1]^d$ and let $\Omega_\mu := [-\mu, \mu]^d$ for a parameter μ.

We denote by $H^q(\Omega)$ the standard periodic Sobolev spaces on Ω and by $H^{q,p}(\Omega; Y)$ periodic Bochner-Sobolev spaces on $\Omega \times Y$. The corresponding norms are $\|\cdot\|_{H^q}$ and $\|\cdot\|_{H^{q,p}}$. Furthermore, $W^{1,q}$ and $W^{p,\infty}$ denote standard Sobolev spaces, $\|\cdot\|_{W^{p,\infty}}$ being the Sobolev supremum norm on Ω: given $\partial_x^\beta u \in L^\infty(\Omega)$ for a multi-index β with $0 \le |\beta| \le p$, it holds that

$$\|u\|_{W^{p,\infty}} = \max_{|\beta| \le p} \sup_{x \in \Omega} |\partial_x^\beta u| \,.$$

Furthermore, we make frequent use of the Sobolev inequality stating that given $u \in H^2(\Omega)$ for dimension $d \le 3$, it holds that

$$(1.3) \qquad\qquad \sup_{x \in \Omega} |u(x)| \le C \|u\|_{H^2(\Omega)} \,.$$

In general, we use capital letters to refer to solutions on the whole domain Ω and for time $[0, T]$, $\mathbf{M} : \Omega \times [0, T] \to \mathbb{R}^3$. When instead considering a micro problem set on Ω_μ, we use lower case letters to denote the solution. By $\boldsymbol{\nabla}\mathbf{M}$ we denote the Jacobian matrix of the vector-valued function $\mathbf{M}$. We assume in general that scalar and cross product between a vector-valued and a matrix-valued function are done column-wise, while the divergence operator is applied row-wise.

The differential operator L is defined such that for $u \in H^2(\Omega)$,

$$(1.4) \qquad Lu(x) = \nabla \cdot (a^\varepsilon(x)\nabla u(x)),$$

where a^ε is a highly oscillatory, smooth, scalar coefficient function. Moreover, we denote by $\mathcal{L}$ the corresponding operator acting on vector-valued functions in $\mathbb{R}^3$,

$$\mathcal{L}\mathbf{m} = \begin{bmatrix} Lm^{(1)} & Lm^{(2)} & Lm^{(3)} \end{bmatrix}^T,$$

where $m^{(i)}$ is the i-th component of $\mathbf{m}$.

1.2. Homogenized equation for the periodic problem. The homogenization of (1.1) with a periodic material coefficient was studied in [17]. Using the setup considered there makes it possible to obtain bounds for the errors introduced when using HMM. It therefore is the scenario that we focus on in the subsequent proofs. Specifically, we assume that $\mathbf{M}^\varepsilon$ satisfies (1.1) with $a^\varepsilon = a(x/\varepsilon)$ on a domain $\Omega = [0, \ell]^d \subset \mathbb{R}^d$, where $d = 1, 2$ or 3 and $\ell > 0$, and for $0 \leq t \leq T$. Then it is shown in [17] that the corresponding homogenized equation is

$$(1.5a) \qquad \partial_t \mathbf{M}_0 = -\mathbf{M}_0 \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{M}_0 \mathbf{A}^H) - \alpha \mathbf{M}_0 \times \mathbf{M}_0 \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{M}_0 \mathbf{A}^H),$$

$$(1.5b) \qquad \mathbf{M}_0(x, 0) = \mathbf{M}_{\text{init}}(x),$$

for $0 \leq t \leq T$ and $x \in \Omega$. The constant homogenized coefficient matrix $\mathbf{A}^H \in \mathbb{R}^{d \times d}$ is as in elliptic homogenization theory given by

$$(1.6) \qquad \mathbf{A}^H := \int_Y a(y) \left(\mathbf{I} + (\boldsymbol{\nabla}_y \boldsymbol{\chi})^T \right) dy,$$

where $\boldsymbol{\chi} : \mathbb{R}^d \to \mathbb{R}^d$ is the solution to the elliptic cell problem

$$(1.7) \qquad \boldsymbol{\nabla}_y \cdot (a(y)\boldsymbol{\nabla}_y \boldsymbol{\chi}(y)) = -\nabla_y a(y).$$

Note that (1.7) determines $\boldsymbol{\chi}$ only up to a constant. Throughout this article, we assume that this constant is chosen such that $\boldsymbol{\chi}$ has zero average.

For the difference between $\mathbf{M}^\varepsilon$ and $\mathbf{M}_0$, the following result was proved in [17].

THEOREM 1.1. *Assume that* $\mathbf{M}^\varepsilon \in C^1([0, T]; H^2(\Omega))$ *is a classical solution to* (1.1) *with a periodic material coefficient* $a^\varepsilon = a(x/\varepsilon)$ *where* $a \in C^\infty(\Omega)$ *and that* $a_{\min} \leq a(x) \leq a_{\max}$ *for some constants* $a_{\min}, a_{\max} > 0$. *Assume that there is a constant* M *independent of* ε *such that* $\|\boldsymbol{\nabla}\mathbf{M}^\varepsilon(\cdot, t)\|_{L^\infty} \leq M$ *for* $0 \leq t \leq T$. *Moreover, suppose that* $\mathbf{M}_0 \in C^\infty([0, T]; H^\infty(\Omega))$ *is a classical solution to* (1.5). *We then have*

$$(1.8) \qquad \|\mathbf{M}^\varepsilon(\cdot, t) - \mathbf{M}_0(\cdot, t)\|_{L^2} \leq C\varepsilon, \qquad 0 \leq t \leq T,$$

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

Remark 1.2. It is easy to check that there is a constant C independent of ε such that $\|\boldsymbol{\nabla}\mathbf{M}^\varepsilon\|_{L^2} \leq C$. The assumption here is that this also holds for L^∞, as ample numerical evidence suggests. It also follows from the results in [17] that such an L^∞-bound is true for the homogenized solution with higher order correction terms.

2. Heterogeneous multiscale methods. The concept of heterogeneous multiscale methods was first introduced by E and Engquist in [21]. It provides a general approach to treat multiscale problems with scale separation, where a description of the microscopic problem is available but would be too computationally expensive to use throughout the whole domain. The idea is therefore to use numerical homogenization with the goal to get a good approximation to the effective solution of the original problem. In general, HMM models involve three parts:

1. *Macro model*: an incomplete model for the whole computational domain, discretized with a coarse grid, that is set up in such a way that some data is missing.
2. *Micro model*: an exact model discretized with a grid resolving the fine ε-scale, which is however only solved on a small domain, where it is feasible to use the expensive description.
3. *Upscaling*: an averaging procedure that uses the data obtained when solving the micro problem to generate the quantity needed to complete the macro model.

It is important to make sure that micro and macro model are consistent, which is typically achieved by choosing the initial data for the micro problem as a restriction of the current macro solution.

The HMM framework has been successfully applied to a wide range of applications; see for instance the surveys in [22, 1]. In this paper, we aim to find a good way to set up HMM for the Landau Lifshitz equation (1.1).

In general, the error in the HMM solution consists of two major components apart from discretization errors: an error term related to the fact that the solution to the effective equation is approximated instead of the original one and the so-called HMM error. The HMM error in turn depends on the upscaling error, the error introduced in the data estimation process [1]. In case of the Landau-Lifshitz equation with a periodic material coefficient, an estimate for the first error term is given by Theorem 1.1. The L^2-homogenization error is $\mathcal{O}(\varepsilon)$. In this paper, we focus on estimates for the upscaling error and investigate how it is influenced by different choices of HMM-models.

We consider three different setups. All three are based on the same micro model, the full Landau-Lifshitz equation,

$$
\text{(2.1a)} \qquad \partial_t \mathbf{m}^\varepsilon = -\mathbf{m}^\varepsilon \times \mathcal{L}\mathbf{m}^\varepsilon - \alpha \mathbf{m}^\varepsilon \times \mathbf{m}^\varepsilon \times \mathcal{L}\mathbf{m}^\varepsilon, \quad x \in \Omega,\ 0 < t \le \eta,
$$
$$
\text{(2.1b)} \qquad \mathbf{m}^\varepsilon(0, x) = \mathbf{m}_{\text{init}}(x) = \mathcal{R}(\mathbf{M}),
$$

with periodic boundary conditions. The initial data $\mathbf{m}_{\text{init}}$ is assumed to be a restriction $\mathcal{R}$ of the macro data $\mathbf{M}$ such that the micro and macro model are consistent and $|\mathbf{m}_{\text{init}}| \equiv 1$. One possible choice to obtain such initial data is by using a normalized interpolation polynomial based on the macro data. In this paper, we assume a solution to the micro problem in the whole domain Ω with periodic boundary conditions. In practice, one would solve the micro problem only on a small domain $[-\mu', \mu']^d$, where $0 < \mu' \ll \ell$. However, this requires a choice of boundary conditions which introduce some additional error. In [18], a possible way to handle this is discussed. To simplify the following analysis, we avoid dealing with this issue here and assume a solution in the whole domain. For the averaging we then consider only the solution in a box $\Omega_\mu = [-\mu, \mu]^d$ in space and an interval $[0, \eta]$ in time, where $\mu \sim \varepsilon$ and $\eta \sim \varepsilon^2$. This matches the scales of the fast variations in the problem as explained in [17]. Note that when choosing the micro domain size proportional to the fast variations, those can be resolved for any given ε using a fixed computational mesh. This implies that with

HMM, the resulting overall computational cost is independent of ε and only depends on the number of grid points chosen to discretize the micro and macro model in space and time. More details about the complexity and computational costs for the HMM method described here can be found in [18].

The other two HMM components, macro model and upscaling, differ between the models. We suppose that the macro models should have the general form of the effective equation (1.5) and consider three different choices of missing data in the model as described in the following.

(M1) *Flux model.* We choose the macro model

$$(2.2a) \qquad \partial_t \mathbf{M} = -\mathbf{M} \times \mathbf{\nabla} \cdot \mathbf{F}_1 - \alpha \mathbf{M} \times \mathbf{M} \times \mathbf{\nabla} \cdot \mathbf{F}_1 \,,$$

$$(2.2b) \qquad \mathbf{M}(x, 0) = \mathbf{M}_{\text{init}}(x) \,,$$

where the missing information to complete the model is the flux $\mathbf{F}_1$. In case of a periodic material coefficient, $\mathbf{F}_1$ would ideally be $\mathbf{\nabla}\mathbf{M}\mathbf{A}^H$. Then (2.2) coincides with (1.5). To obtain $\mathbf{F}_1$, we average the product of the material coefficient and gradient of the solution to the micro problem in space and time using averaging kernels K_μ, K_η^0 as explained in more detail in Section 4,

$$\mathbf{F}_1 = \int_0^\eta \int_{-\mu}^\mu K_\eta^0(t) K_\mu(x) a^\varepsilon \mathbf{\nabla} \mathbf{m}^\varepsilon \, dx dt \,.$$

(M2) *Field model.* Here the macro model is given by

$$\partial_t \mathbf{M} = -\mathbf{M} \times \mathbf{F}_2 - \alpha \mathbf{M} \times \mathbf{M} \times \mathbf{F}_2 \,,$$

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

where $\mathbf{F}_2$ takes the role of the effective field. In the periodic case, $\mathbf{F}_2$ should hence approximate $\mathbf{\nabla} \cdot (\mathbf{\nabla}\mathbf{M}\mathbf{A}^H)$. In general, $\mathbf{F}_2$ is defined as the average of the operator $\mathcal{L}$ applied to the solution to the micro problem,

$$\mathbf{F}_2 = \int_0^\eta \int_{-\mu}^\mu K_\eta^0(t) K_\mu(x) \mathbf{\nabla} \cdot (a^\varepsilon \mathbf{\nabla} \mathbf{m}^\varepsilon) \, dx dt \,.$$

(M3) *Torque model.* The third macro model we consider is

$$\partial_t \mathbf{M} = -\mathbf{F}_3 - \alpha \mathbf{M} \times \mathbf{F}_3 \,,$$

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

which means that for a periodic material coefficient, $\mathbf{F}_3$ should approximate the torque $\mathbf{M} \times \mathbf{\nabla} \cdot (\mathbf{\nabla}\mathbf{M}\mathbf{A}^H)$. Here $\mathbf{F}_3$ is given by

$$\mathbf{F}_3 = \int_0^\eta \int_{-\mu}^\mu K_\eta^0(t) K_\mu(x) \mathbf{m}^\varepsilon \times \mathbf{\nabla} \cdot (a^\varepsilon \mathbf{\nabla} \mathbf{m}^\varepsilon) \, dx dt \,.$$

In the following, we prove estimates for the upscaling error in each of the three models, (M1) - (M3), when a^ε is periodic,

$$E_1 := \left| \mathbf{F}_1 - \mathbf{\nabla}\mathbf{M}_0\mathbf{A}^H \right| \,,$$

$$E_2 := \left| \mathbf{F}_2 - \mathbf{\nabla} \cdot (\mathbf{\nabla}\mathbf{M}_0\mathbf{A}^H) \right| \,,$$

$$E_3 := \left| \mathbf{F}_3 - \mathbf{M}_0 \times (\mathbf{\nabla} \cdot (\mathbf{\nabla}\mathbf{M}_0\mathbf{A}^H)) \right| \,,$$

under the assumption that $\partial_x^\beta \mathbf{m}_{\text{init}}(0, 0) = \partial_x^\beta \mathbf{M}_0(x_\mathrm{M}, t_\mathrm{M})$ for multi-indices β with $|\beta| \le 2$, where $(x_\mathrm{M}, t_\mathrm{M})$ is the macro point we average around.

3. Homogenized solution and correctors for a periodic micro problem.
To be able to prove estimates for the upscaling errors E_1 - E_3 given a periodic material coefficient, we make use of the estimates for the error between the actual and the homogenized solution to (2.1) as well as the corresponding corrected approximations that were derived in [17].

Let $\mathbf{m}^\varepsilon$ be the solution to (2.1) given $a^\varepsilon(x) = a(x/\varepsilon)$. Then, according to [17], the corresponding homogenized solution is $\mathbf{m}_0(x,t)$, which for $0 \leq t \leq \eta$ satisfies

(3.1a) $\qquad \partial_t \mathbf{m}_0 = -\mathbf{m}_0 \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{m}_0 \mathbf{A}^H) - \alpha \mathbf{m}_0 \times \mathbf{m}_0 \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{m}_0 \mathbf{A}^H), \quad x \in \Omega,$

(3.1b) $\ \mathbf{m}_0(x,0) = \mathbf{m}_{\mathrm{init}}(x),$

with periodic boundary conditions and where $\mathbf{m}_{\mathrm{init}}$ is chosen as in (2.1).

We consider a short time interval $[0, T^\varepsilon]$, with an ε-dependent final time

(3.2) $$T^\varepsilon := \varepsilon^\sigma T, \qquad 1 < \sigma \leq 2.$$

This is still sufficiently long time for the HMM micro problems with final time $\eta \sim \varepsilon^2$. In [17], it is shown that for such a time interval, one obtains improved approximations to the solution to (2.1), $\mathbf{m}^\varepsilon(x)$, when not only considering $\mathbf{m}_0$ but a truncated asymptotic expansion

(3.3) $$\tilde{\mathbf{m}}_J^\varepsilon(x,t) = \mathbf{m}_0(x,t) + \sum_{j=1}^J \varepsilon^j \mathbf{m}_j \left(x, \frac{x}{\varepsilon}, t, \frac{t}{\varepsilon^2} \right), \qquad J > 0,$$

where the correctors $\mathbf{m}_j(x,y,t,\tau)$, $j > 0$ satisfy linear differential equations in the fast variables y and τ,

$$\partial_\tau \mathbf{m}_j = -\mathbf{m}_0 \times \mathcal{L}_{yy} \mathbf{m}_j - \alpha \mathbf{m}_0 \times \mathbf{m}_0 \times \mathcal{L}_{yy} \mathbf{m}_j - \mathbf{f}_j,$$
$$\mathbf{m}_j(x,y,t,0) = 0.$$

Here the forcing $\mathbf{f}_j$ depends only on the lower order terms $\mathbf{m}_k$, $0 \leq k < j$. The operator $\mathcal{L}_{yy}$ is the vector-equivalent to L_{yy}, which is defined such that for $u(x,y) \in H^{0,2}(\Omega; Y)$,

(3.5) $$L_{yy} u(x,y) = \nabla_y \cdot (a(y) \nabla_y u(x,y)).$$

As explained in [17], the form of the first corrector $\mathbf{m}_1$ is

(3.6) $$\mathbf{m}_1(x,y,t,\tau) = \boldsymbol{\nabla} \mathbf{m}_0(x,t) \boldsymbol{\chi}(y) + \mathbf{v}(x,y,t,\tau),$$

where $\boldsymbol{\chi}$ is the solution to (1.7) and $\mathbf{v}$ is an oscillatory, decaying term that is discussed in Section 3.2.

3.1. Energy and error estimates. For convenience of the reader, we here give a summary of the estimates from [17] that are most crucial for the derivations in this paper. In contrast to [17], we here require higher regularity of $\mathbf{m}_0$ for reasons of simplicity. Otherwise we use the same assumptions. In particular, we assume that
  (A1) the material coefficient $a \in C^\infty(\Omega)$ is a periodic function such that there are positive constants $a_{\min}, a_{\max} > 0$ satisfying $a_{\min} < a(x) < a_{\max}$.
  (A2) the initial data function is normalized, $|\mathbf{m}_{\mathrm{init}}(x)| = 1$, which implies that $|\mathbf{m}^\varepsilon(x,t)| = |\mathbf{m}_0(x,t)| = 1$ for any $t \geq 0$. From this property it follows that given a multi-index β with $|\beta| = 1$,

$$0 = \partial^\beta |\mathbf{m}_0|^2 = 2\mathbf{m}_0 \cdot \partial^\beta \mathbf{m}_0,$$

  and thus $\mathbf{m}_0$ and $\boldsymbol{\nabla} \mathbf{m}_0$ are orthogonal.

(A3) the damping coefficient α is positive and it holds that $0 < \alpha \le 1$. Moreover, $\varepsilon = \ell/n$ for some $n \in \mathbb{N} \gg 1$, which implies $0 < \varepsilon \ll 1$.

(A4) $\mathbf{m}^\varepsilon \in C^1([0, T^\varepsilon]; H^{s+1}(\Omega))$ with $s \ge 1$ is a classical solution to (2.1) and there is a constant M independent of ε such that

$$\|\boldsymbol{\nabla}\mathbf{m}^\varepsilon(\cdot, t)\|_{L^\infty} \le M, \qquad 0 \le t \le T^\varepsilon.$$

(A5) $\mathbf{m}_0 \in C^\infty(0, T; H^\infty(\Omega))$ is a classical solution to (3.1).

As shown in [17], it then holds for any $q \ge 0$ that the $H^{q,\infty}$-norms of the first two correctors, $\mathbf{m}_1$ and $\mathbf{m}_2$, are bounded uniformly in the fast time variable τ, while the norms of higher order correctors grow algebraically with τ. Specifically, it holds for all $p, q \ge 0$ and $0 \le t \le T^\varepsilon$ that

$$(3.7) \qquad \|\mathbf{m}_j(\cdot, \cdot, t, t/\varepsilon^2)\|_{H^{q,p}} \le C \begin{cases} 1, & j = 1, 2, \\ \varepsilon^{(\sigma-2)(j-2)}, & j \ge 3, \end{cases}$$

where the constant C depends on T but is independent of ε. For the approximating $\tilde{\mathbf{m}}_J^\varepsilon$ in (3.3), it holds for $0 \le t \le T^\varepsilon$ that

$$(3.8) \qquad \|\tilde{\mathbf{m}}_J^\varepsilon\|_{W^{q,\infty}} \le C\varepsilon^{\min(0, 1-q)}, \qquad q \ge 0.$$

Moreover, consider the error introduced when approximating $\mathbf{m}^\varepsilon$ by $\tilde{\mathbf{m}}_J^\varepsilon$. Under the given assumptions, it holds for $0 \le t \le T^\varepsilon$ and $q \le s$ that

$$(3.9) \qquad \|\mathbf{m}^\varepsilon(\cdot, t) - \tilde{\mathbf{m}}_J^\varepsilon(\cdot, t)\|_{H^q} \le C\varepsilon^{2-q+(\sigma-1)(J-1)}, \qquad J \ge 1,$$

where the constant C is independent of ε but depends on M in (A4) and T. This estimate shows that the approximations improve with increasing J on the considered time interval.

3.2. The correction term v. In [17], it was shown that both $\mathbf{m}_1$ and the correction term $\mathbf{v}$, which is part of $\mathbf{m}_1$ as given in (3.6), are orthogonal to $\mathbf{m}_0$,

$$(3.10) \qquad \mathbf{v} \perp \mathbf{m}_0, \qquad \mathbf{m}_1 \perp \mathbf{m}_0.$$

Moreover, it was proved that given (A1)-(A5), there are constants $\gamma > 0$ and C independent of ε such that

$$(3.11) \qquad \|\partial_t^k \mathbf{v}(\cdot, \cdot, t, \tau)\|_{H^{p,q}} \le C\varepsilon^{-\gamma\tau}, \qquad k, p, q \ge 0.$$

For the analysis in this paper, an explicit formulation for $\mathbf{v}(x, y, t, \tau)$ is required. To obtain such a description, we use the linear equation that was derived in [17],

$$(3.12\text{a}) \qquad \partial_\tau \mathbf{v} = -\mathbf{m}_0 \times \mathcal{L}_{yy}\mathbf{v} - \alpha\mathbf{m}_0 \times \mathbf{m}_0 \times \mathcal{L}_{yy}\mathbf{v},$$
$$(3.12\text{b}) \qquad \mathbf{v}(x, y, t, 0) = -\boldsymbol{\nabla}\mathbf{m}_0(x, t)\boldsymbol{\chi}(y).$$

We now introduce a lemma that shows a connection between differential equations of the same type as (3.12) to a system of parabolic equations that become Schrödinger equations as $\alpha \to 0$. Then we go on and use that result to derive an explicit solution to (3.12) in terms of the eigenfunctions of the operator $-L_{yy}$.

LEMMA 3.1. *Suppose* $\mathbf{f} \in H^2(Y, \mathbb{R}^3)$ *and* $\mathbf{b} \in \mathbb{R}^3$ *is a given constant vector with* $|\mathbf{b}| = 1$. *Then the solution* $\mathbf{w}$ *to*

$$
\begin{aligned}
(3.13) \qquad \partial_t \mathbf{w}(y,t) &= -\mathbf{b} \times \mathcal{L}_{yy}\mathbf{w}(y,t) - \alpha \mathbf{b} \times \mathbf{b} \times \mathcal{L}_{yy}\mathbf{w}(y,t), \qquad y \in Y, \ t > 0, \\
\mathbf{w}(y,0) &= \mathbf{f}(y),
\end{aligned}
$$

with periodic boundary conditions is given by

$$
(3.14) \qquad \mathbf{w}(y,t) := \mathbf{b}\mathbf{b}^T \mathbf{f}(y) + (\mathbf{I} - \mathbf{b}\mathbf{b}^T)\mathcal{R}e(\mathbf{u}(y,t)) + \mathbf{b} \times \mathcal{I}m(\mathbf{u}(y,t)),
$$

where $\mathbf{u} \in C^1(0,T; H^2(Y; \mathbb{C}^3))$ *solves*

$$
\begin{aligned}
(3.15) \qquad \partial_t \mathbf{u}(y,t) &= -(i - \alpha)\mathcal{L}_{yy}\mathbf{u}(y,t), \qquad y \in Y, \ t > 0, \\
\mathbf{u}(y,0) &= \mathbf{f}(y)
\end{aligned}
$$

with periodic boundary conditions.

Proof. As $\mathbf{u}(y,0)$ is real, it follows immediately that $\mathbf{w}$ given by (3.14) satisfies the initial condition in (3.13). Moreover, since $\mathbf{b}$ is constant,

$$
\begin{aligned}
\mathbf{b} \times \mathcal{L}_{yy}\mathbf{w} &= \mathbf{b} \times [\mathbf{b}\mathbf{b}^T \mathcal{L}_{yy}\mathbf{f} + (\mathbf{I} - \mathbf{b}\mathbf{b}^T)\mathcal{L}_{yy}\mathcal{R}e(\mathbf{u}) + \mathbf{b} \times \mathcal{L}_{yy}\mathcal{I}m(\mathbf{u})] \\
&= \mathbf{b} \times \mathcal{R}e(\mathcal{L}_{yy}\mathbf{u}) + \mathbf{b} \times \mathbf{b} \times \mathcal{I}m(\mathcal{L}_{yy}\mathbf{u}) \\
&= \mathbf{b} \times \mathcal{R}e(\mathcal{L}_{yy}\mathbf{u}) - (\mathbf{I} - \mathbf{b}\mathbf{b}^T)\mathcal{I}m(\mathcal{L}_{yy}\mathbf{u}),
\end{aligned}
$$

where we used the vector triple product identity for the last step. It then follows that

$$
\mathbf{b} \times \mathbf{b} \times \mathcal{L}_{yy}\mathbf{w} = \mathbf{b} \times \mathbf{b} \times \mathcal{R}e(\mathcal{L}_{yy}\mathbf{u}) - \mathbf{b} \times \mathcal{I}m(\mathcal{L}_{yy}\mathbf{u}).
$$

It thus holds that

$$
\begin{aligned}
\mathbf{b} \times \mathcal{L}_{yy}\mathbf{w} &+ \alpha \mathbf{b} \times \mathbf{b} \times \mathcal{L}_{yy}\mathbf{w} \\
&= \mathbf{b} \times [\mathcal{R}e(\mathcal{L}_{yy}\mathbf{u}) - \alpha \mathcal{I}m(\mathcal{L}_{yy}\mathbf{u})] + \mathbf{b} \times \mathbf{b} \times [\mathcal{I}m(\mathcal{L}_{yy}\mathbf{u}) + \alpha \mathcal{R}e(\mathcal{L}_{yy}u)] \\
&= \mathbf{b} \times \mathcal{I}m[(i - \alpha)\mathcal{L}_{yy}\mathbf{u}] - \mathbf{b} \times \mathbf{b} \times \mathcal{R}e[(i - \alpha)\mathcal{L}_{yy}\mathbf{u}].
\end{aligned}
$$

Using (3.15) and exploiting the facts that $\mathbf{b}$ is constant and $\mathbf{f}$ is independent of time, we obtain

$$
\begin{aligned}
\mathbf{b} \times \mathcal{L}_{yy}\mathbf{w} + \alpha \mathbf{b} \times \mathbf{b} \times \mathcal{L}_{yy}\mathbf{w} &= -\mathbf{b} \times \mathcal{I}m(\partial_t \mathbf{u}) + \mathbf{b} \times \mathbf{b} \times \mathcal{R}e(\partial_t \mathbf{u}) \\
&= -\partial_t[(\mathbf{I} - \mathbf{b}\mathbf{b}^T)(\mathcal{R}e(\mathbf{u})) + \mathbf{b} \times \mathcal{I}m(\mathbf{u})] = -\partial_t \mathbf{w},
\end{aligned}
$$

which shows that $\mathbf{w}$ given by (3.14) satisfies (3.13). $\square$

In the following, let $\phi_j(y), \omega_j$ be the eigenfunctions and eigenvalues of the operator $-L_{yy}$, where L_{yy} is given by (3.5), on Y with periodic boundary conditions,

$$
-L_{yy}\phi_j(y) = \omega_j \phi_j(y) \quad \text{on } Y.
$$

As $-L_{yy}$ is a periodic elliptic operator it holds according to standard theory that its eigenvalues are strictly positive and bounded away from zero except for the first eigenvalue, ω_0, which is zero [15],

$$
\omega_0 = 0, \qquad 0 < \omega_j, \quad \text{for} \quad j > 0.
$$

Moreover, the eigenfunctions ϕ_j form an orthonormal basis for $L^2(Y)$. In particular, the first eigenfunction, corresponding to ω_0, is the constant function $\phi_0 \equiv 1$. The eigenfunctions can be chosen to be real, which they are assumed to be in the following. We then obtain the following expression for the correction term $\mathbf{v}$.

LEMMA 3.2. *Let* $\mathbf{f_v}(x,t) := \boldsymbol{\nabla}\mathbf{m}_0(x,t) - i\mathbf{m}_0(x,t) \times \boldsymbol{\nabla}\mathbf{m}_0(x,t)$ *and*

$$(3.16) \qquad \Psi(y,\tau) := \sum_{j=1}^{\infty} \boldsymbol{\chi}_j e^{(-\alpha+i)\omega_j\tau}\phi_j(y),$$

where ϕ_j *and* ω_j, $0 \leq j$, *are the eigenfunctions and eigenvalues of* $-L_{yy}$ *and* $\boldsymbol{\chi}_j$ *are expansion coefficients such that* $\boldsymbol{\chi}(y) = \sum_j \boldsymbol{\chi}_j \phi_j(y)$. *Then*

$$\mathbf{v}(x,y,t,\tau) = -\mathrm{Re}(\mathbf{f_v}(x,t)\Psi(y,\tau))$$

solves (3.12).

Proof. Since $\mathbf{v}$ satisfies a linear differential equation in the fast variables, we write $\mathbf{v}(y,\tau)$ and suppress the dependence on the slow variables, x and t, in the notation throughout this proof. Note also that with respect to the fast variables only, $\mathbf{m}_0$ and $\boldsymbol{\nabla}\mathbf{m}_0$ are constant.

By Lemma 3.1 and using the fact that by (A2), $\mathbf{m}_0$ and $\boldsymbol{\nabla}\mathbf{m}_0$ are orthogonal to each other, we find that the solution to (3.12) is

$$(3.17) \qquad \mathbf{v}(y,\tau) = (\mathbf{I} - \mathbf{m}_0\mathbf{m}_0^T)\mathcal{R}e(\mathbf{u}(y,\tau)) + \mathbf{m}_0 \times \mathcal{I}m(\mathbf{u}(y,\tau)),$$

where $\mathbf{u}$ is the solution to

$$(3.18) \qquad \begin{aligned} \partial_\tau \mathbf{u}(y,\tau) &= -(i-\alpha)\mathcal{L}\mathbf{u}(y,\tau), \qquad y \in Y, \tau > 0, \\ \mathbf{u}(y,0) &= -\boldsymbol{\nabla}\mathbf{m}_0\boldsymbol{\chi}(y). \end{aligned}$$

As (3.18) is a system of three decoupled equations, we can consider each equation separately and solve it in terms of the eigenfunctions of $-L_{yy}$. Let $u(y,\tau)$ denote the first component in $\mathbf{u}$. Then we can define $u_j(\tau)$ such that

$$u(y,\tau) = \sum_{j=0}^{\infty} u_j(\tau)\phi_j(y).$$

Note that $\boldsymbol{\chi}_0 = 0$ as $\boldsymbol{\chi}$ has zero average by definition. By (3.18) and the orthogonality of the eigenfunctions, we deduce that

$$\partial_\tau u_j(\tau) = (i-\alpha)\omega_j u_j(\tau), \qquad u_j(0) = -\nabla m_0^{(1)} \cdot \boldsymbol{\chi}_j,$$

and consequently,

$$u_j(\tau) = -\nabla m_0^{(1)} \cdot \boldsymbol{\chi}_j e^{(i-\alpha)\omega_j\tau}.$$

For the second and third components in $\mathbf{u}$, we obtain the same result but with initial conditions involving $\nabla m_0^{(2)}$ and $\nabla m_0^{(3)}$, respectively. Hence, in total it holds that

$$\mathbf{u}(y,\tau) = -\boldsymbol{\nabla}\mathbf{m}_0 \sum_{j=1}^{\infty} \boldsymbol{\chi}_j e^{(i-\alpha)\omega_j\tau}\phi_j(y) = -\boldsymbol{\nabla}\mathbf{m}_0\Psi(y,\tau),$$

where $\Psi(y,\tau)$ is defined as in (3.16). Putting this explicit expression for $\mathbf{u}$ into (3.17), then results in

$$\begin{aligned} \mathbf{v}(y,\tau) &= -\boldsymbol{\nabla}\mathbf{m}_0\mathcal{R}e(\Psi(y,\tau)) - \mathbf{m}_0 \times \boldsymbol{\nabla}\mathbf{m}_0\mathcal{I}m(\Psi(y,\tau)) \\ &= -\mathcal{R}e(\boldsymbol{\nabla}\mathbf{m}_0\Psi(y,\tau) - i\mathbf{m}_0 \times \boldsymbol{\nabla}\mathbf{m}_0\Psi(y,\tau)). \end{aligned}$$

This completes the proof. $\qquad\qquad\Box$

To gain a more intuitive understanding, note that $\mathbf{v}(y,\tau)$ can also be written as

$$(3.19) \qquad \mathbf{v}(y,\tau) = -\sum_{j=1}^{\infty} \chi_j \phi_j(y) e^{-\alpha \omega_j \tau} [\cos(\omega_j \tau) \boldsymbol{\nabla} \mathbf{m}_0 + \sin(\omega_j \tau) \mathbf{m}_0 \times \boldsymbol{\nabla} \mathbf{m}_0].$$

As $\boldsymbol{\nabla}\mathbf{m}_0$ and $\mathbf{m}_0 \times \boldsymbol{\nabla}\mathbf{m}_0$ are orthogonal to $\mathbf{m}_0$ and each other, this clearly shows that $\mathbf{v}$ lies in the subspace orthogonal to $\mathbf{m}_0$ and can be written in terms of two orthogonal vectors spanning this subspace multiplied by coefficients that oscillate with τ. For $\alpha > 0$, all the components of $\mathbf{v}$ are damped away with increasing τ, with stronger damping for higher modes. Note that the sum in (3.19) starts from $j = 1$. There is no contribution from the constant mode, indicating that $\mathbf{v}$ has zero average.

4. Averaging. In order to get a good approximation of the missing quantity for the macro model in our HMM scheme, it is crucial to have efficient averaging techniques that allow us to control how fast the averaged micro model data converges to the required effective quantity. To achieve this, one can use smooth, compactly supported averaging kernels as introduced in [12], [5].

DEFINITION 4.1 ([5]). *A function K is in the space of smoothing kernels $\mathbb{K}^{p,q}$ if*
1. *$K^{(q+1)} \in BV(\mathbb{R})$ and K has compact support in $[-1,1]$, $K \in C_c^q([-1,1])$.*
2. *K has p vanishing moments,*

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

Typically, we do not want to average over $[-1,1]$ but over small boxes of size proportional to ε or ε^2. For this purpose, let $K_\mu(x)$ denote a scaled version of $K(x)$,

$$K_\mu(x) = \frac{1}{\mu} K\left(\frac{x}{\mu}\right).$$

Moreover, when considering problems in d space dimensions with $d > 1$, $K_\mu(x)$ is to be understood as

$$K_\mu(x) = K_\mu(x_1) \cdots K_\mu(x_d).$$

Note that as K has compact support and $K^{(q+1)} \in BV(\mathbb{R})$, it holds that $K \in W^{1,q+1}(\mathbb{R})$ and $K \in L^2(\mathbb{R})$.

4.1. Kernels K^0. Often, the averaging kernels used for HMM are chosen to be symmetric around zero and have nonzero-values almost everywhere in $[-1,1]$. However, for our application it is advantageous to do time averaging such that we obtain an approximation for the effective quantity at time $t = 0$ based only on the values of the microscopic solution for $t \ge 0$. As the subsequent proofs require kernels $K \in \mathbb{K}^{p,q}$, we therefore show that $\mathbb{K}^{p,q}$ contains a subspace $\mathbb{K}_0^{p,q}$ such that $K^0(t) = 0$ for $t \le 0$ when $K^0 \in \mathbb{K}_0^{p,q}$. To construct such kernels, consider the ansatz

$$(4.1) \qquad K^0(t) = \begin{cases} t^{q+1}(1-t)^{q+1} P(t), & 0 < t < 1 \\ 0, & otherwise, \end{cases}$$

where P is a polynomial in $\mathbb{P}^p$, the space of of polynomials of degree p,

$$P(t) = c_0 + c_1 t + \dots + c_p t^p.$$

As explained in [11], it is beneficial to choose this type of ansatz since it typically results in better numerical stability compared to an approach where the coefficients of K^0 are computed directly.

One can easily see that due to the term $t^{q+1}(1-t)^{q+1}$, the first q derivatives of K^0 as given by (4.1) vanish at zero and one, which together with continuity implies that the first requirement in Definition 4.1 is satisfied.

To show that there indeed exists a unique polynomial P in $\mathbb{P}^p$ such that K^0 as given in (4.1) also satisfies the second requirement in Definition 4.1 and hence is in $\mathbb{K}^{p,q}$, we define the weighted inner product $\langle \cdot, \cdot \rangle_w$ by

$$\langle u, v \rangle_w := \int_0^1 u(t)v(t)t^{q+1}(1-t)^{q+1}dt, \qquad \|u\|_w := \langle u, u \rangle_w.$$

This allows us to rewrite the second condition in Definition 4.1 as

(4.2)
$$\langle P, 1 \rangle_w = 1,$$
$$\langle P, t^k \rangle_w = 0, \quad 1 \le k \le p.$$

Let now ϕ_j, $j = 0, \dots, p$ be orthogonal polynomials with respect to $\langle \cdot, \cdot \rangle_w$, satisfying the recurrence formula

$$\phi_0 = 1, \quad \phi_1 = (x - \alpha_0)\phi_0, \quad \phi_{k+1} = (x - \alpha_k)\phi_k - \beta_k \phi_{k-1},$$

where $\alpha_k = \frac{\langle \phi_k, x\phi_k \rangle_w}{\|\phi_k\|_w^2}$ and $\beta_k = \frac{\|\phi_k\|_w^2}{\|\phi_{k-1}\|_w^2}$. Then it holds that $\phi_j \in \mathbb{P}^j$ and together the $\phi_j, j = 0, \dots, p$ form an orthogonal basis for $\mathbb{P}^p$. We can hence expand

(4.3)
$$P(t) = \sum_{j=1}^p p_j \phi_j(t), \qquad t^k = \sum_{j=0}^p c_{jk}\phi_j(t),$$

where the coefficients c_{jk} are uniquely determined [19]. In particular, $c_{jj} = 1$ and $c_{jk} = 0$ for $k > j$. Expressing the inner product in (4.2) in terms of the expansions (4.3) yields

$$\langle P, t^k \rangle_w = \sum_{j=1}^k p_j c_{jk} \|\phi_j\|_w^2,$$

which implies that (4.2) is satisfied when the coefficients p_j are the solution to

$$\begin{bmatrix} \|\phi_0\|_w^2 & 0 & 0 & \dots & 0 \\ c_{01}\|\phi_0\|_w^2 & \|\phi_1\|_w^2 & 0 & \dots & 0 \\ c_{02}\|\phi_0\|_w^2 & c_{12}\|\phi_1\|_w^2 & \|\phi_2\|_w^2 & \dots & 0 \\ \vdots & \vdots & & \ddots & \\ c_{0p}\|\phi_0\|_w^2 & c_{1p}\|\phi_1\|_w^2 & \dots & & \|\phi_p\|^2 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \\ p_p \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since the matrix here is triangular with strictly positive diagonal elements, the system has a unique solution, which proves that there exists a unique polynomial P of degree at most p such that $K^0 \in \mathbb{K}_0^{q,p} \subset \mathbb{K}^{q,p}$.

Remark 4.2. In practice, there is a quicker way to determine the coefficients c_j of the polynomial $P(t)$. Let $I_j = \int_0^1 t^{q+1+j}(1-t)^{q+1}dt$, then it has to hold that the

vector containing the coefficients c_j solves the linear system

$$
\begin{bmatrix}
I_0 & I_1 & \ldots & I_p \\
I_1 & I_2 & \ldots & I_{p+1} \\
\vdots & & \ddots & \\
I_r & I_{p+1} & \ldots & I_{2p}
\end{bmatrix}
\begin{bmatrix}
c_0 \\
c_1 \\
\vdots \\
c_p
\end{bmatrix}
=
\begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix}.
$$

4.2. Averaging in space. The following lemma from [5] gives a precise convergence rate in terms of ε/η when averaging a purely periodic function f with a kernel over a one-dimensional interval. By choosing a kernel with high regularity, one can achieve very fast convergence to the corresponding average.

LEMMA 4.3 ([5]). *Let $f : \mathbb{R} \to \mathbb{R}$ be a 1-periodic continuous function, and let $K \in \mathbb{K}^{p,q}$. Then, with $\bar{f} = \int_0^1 f(y)dy$,*

$$
\left| \int_{\mathbb{R}} K_\mu(x) f\left(\frac{x}{\varepsilon}\right) dx - \bar{f} \right| \leq C |f|_\infty \left(\frac{\varepsilon}{\mu}\right)^{q+2}
$$

and when $r \in \mathbb{Z}^+$,

$$
\left| \int_{\mathbb{R}} K_\mu(x) x^r f\left(\frac{x}{\varepsilon}\right) dx \right| \leq C
\begin{cases}
|f|_\infty \left(\frac{\varepsilon}{\mu}\right)^{q+2} \mu^r, & 1 \leq r \leq p, \\
|f|_\infty \left(\frac{\varepsilon}{\mu}\right)^{q+2} \mu^r + |\bar{f}| \mu^r, & r > p,
\end{cases}
$$

where the constant C is independent of ε, μ, f or x but may depend on K, p, q and r.

In [5], this lemma is proved for a continuous f since all derivatives involved in the proof are treated in the classical sense. However, when the derivatives are seen in a weak sense, the lemma also applies to $f \in L^\infty$, as explained in [6].

In [6], an averaging lemma for functions $f(x,y)$ that only are 1-periodic in the second variable is derived. In the following, we give a variation of that lemma which is adapted for Bochner-Sobolev spaces and higher dimensions.

LEMMA 4.4. *Let $Y = [0,1]^d$ and $\Omega_\mu = [-\mu,\mu]^d$ for $d \in \mathbb{N}$. Suppose $f(x,y)$ is 1-periodic in y and $\partial_x^\beta f \in L^\infty(\Omega_\mu; L^\infty(Y))$ for $0 \leq |\beta| \leq p+1$ and assume that $K \in \mathbb{K}^{p,q}$. Then, with $\bar{f}(x) := \int_Y f(x,y)dy$,*

$$
\left| \int_{\mathbb{R}^d} K_\mu(x) f(x, x/\varepsilon) dx - \bar{f}(0) \right| \leq C \sup_{y \in Y} \| f(\cdot, y) \|_{W^{p+1,\infty}(\Omega_\mu)} \left(\left(\frac{\varepsilon}{\mu}\right)^{q+2} + \mu^{p+1} \right),
$$

where the constant C does not depend on μ or f but may depend on K, p and q.

Proof. We first assume that $\partial_x^\beta f \in C(\Omega; L^\infty(Y))$ for $0 \leq |\beta| \leq p+1$. Then we obtain via Taylor expansion of $f(x, x/\varepsilon)$ that

$$
\int_{\mathbb{R}^d} K_\mu(x) f(x, x/\varepsilon) dx = \int_{\Omega_\mu} K_\mu(x) f(0, x/\varepsilon) dx
$$

$$
+ \sum_{1 \leq |\beta| \leq p} \frac{1}{\beta!} \int_{\Omega_\mu} K_\mu(x) \partial_x^\beta f(0, x/\varepsilon) x^\beta dx + \sum_{|\beta|=p+1} \int_{\Omega_\mu} K_\mu(x) R_\beta(x) x^\beta dx
$$

$$
=: I + II + III,
$$

where $R_\beta(x)$ is the remainder in integral form,

$$R_\beta(x) = \frac{|\beta|}{\beta!} \int_0^1 (1-z)^p \partial_x^\beta f(zx, x/\varepsilon)\, dz\,.$$

The terms I and II can be bounded using Lemma 4.3. We consider one coordinate direction at a time. For this purpose, assume that we have a multi-index $\beta = [\beta_1, ..., \beta_d]$, coordinates $x = (x_1, ...x_d)$ and let

$$g_0(y_1, ..., y_d) := \partial_x^\beta f(0, y) = \partial_x^\beta f(0, ..., 0, y_1, ..., y_d),$$

$$g_n(y_{n+1}, ..., y_d) := \int_{-\mu}^\mu K_\mu(x_n) g_{n-1}(x_n/\varepsilon, y_{n+1}, ..., y_d) x_n^{\beta_n}\, dx_n, \qquad 1 \le n \le d-1$$

$$g_d := \int_{-\mu}^\mu K_\mu(x_d) g_{d-1}(x_d/\varepsilon) x_d^{\beta_d}\, dx_d = \int_{\Omega_\mu} \mathcal{K}_\mu(x) \partial_x^\beta f(0, x/\varepsilon) x^\beta\, dx,$$

and

$$h_n(x_n) := \int_0^1 \cdots \int_0^1 g_{n-1}(x_n, y_{n+1}, ..., y_d) dy_{n+1} \cdots dy_d.$$

Note first that due to the fact that $\mu < 1$, we obtain by iterative application of Lemma 4.3 that

(4.4)
$$\begin{aligned}
\sup_{y \in Y} |g_n(y_{n+1}, ..., y_d)| &= \sup_{y \in Y} \left| \int_{-\mu}^\mu K_\mu(x_n) g_{n-1}(x_n/\varepsilon, y_{n+1}, ..., y_d) x_n^{\beta_n}\, dx_n \right| \\
&\le C \sup_{y \in Y} |g_{n-1}(y_n, ..., y_d)| \left((\varepsilon/\mu)^{q+2} + \delta_{\beta_n=0} \right) \\
&\le C \sup_{y \in Y} |\partial_x^\beta f(0, y)| \prod_{j=1}^n \left((\varepsilon/\mu)^{q+2} + \delta_{\beta_j=0} \right),
\end{aligned}$$

where $\delta_{\beta_j=0}$ indicates that there is a term of order one when $\beta_j = 0$, an upper bound for the average in coordinate direction j.

In case of I, we have $|\beta| = 0$. An application of Lemma 4.3 then yields that for $1 \le j \le d$,

$$|\bar{g}_j - \bar{g}_{j-1}| = \left| \int_{-\mu}^\mu K_\mu(x_j) h_j(x_j/\varepsilon) dx_j - \int_0^1 h_j(x_j) dx_j \right| \le C \sup_{y_j \in [0,1]} |h_j(y_j)| \left(\frac{\varepsilon}{\mu} \right)^{q+2},$$

and as a consequence of (4.4), it holds that

$$\sup_{y_j \in [0,1]} |h_j(x_j)| \le \sup_{y_j,...,y_d \in [0,1]} |g_{j-1}(y_n, ..., y_d)| \le C \sup_{y \in Y} |f(0, y)|.$$

Using the fact that $I = g_d = \bar{g}_d$ and $\bar{f}(0) = \bar{g}_0$ we hence obtain

(4.5)
$$|I - \bar{f}(0)| \le |I - \bar{g}_{d-1}| + \sum_{j=1}^{d-1} |\bar{g}_j - \bar{g}_{j-1}| \le C \sup_{y \in Y} |f(0, y)| \left(\frac{\varepsilon}{\mu} \right)^{q+2}.$$

To estimate the integrals in II, consider $1 \le |\beta| \le p$. It then follows by (4.4) that

$$\left| \int_{\Omega_\mu} K_\mu(x) \partial_x^\beta f(0, x/\varepsilon) x^\beta\, dx \right| = |g_d| \le C \sup_{y \in Y} |\partial_x^\beta f(0, y)| \left(\frac{\varepsilon}{\mu} \right)^{q+2},$$

where the last step follows since we know that $|\beta| > 0$, there is at least one direction j such that $\beta_j > 0$. Consequently, we obtain

$$(4.6) \qquad |II| \leq C \max_{1 \leq |\beta| \leq p} \sup_{y \in Y} |\partial_x^\beta f(0, y)| \left(\frac{\varepsilon}{\mu}\right)^{q+2}.$$

To bound III we use the fact that

$$\sup_{x \in \Omega_\mu} |R_\beta(x)| \leq C \frac{|\beta|}{\beta!} \sup_{x \in [-1,1]^d} \left| \int_0^1 (1-z)^p \partial_x^\beta f\left(z\mu x, \frac{\mu x}{\varepsilon}\right) dz \right|$$

$$\leq C \sup_{x \in \Omega_\mu} \sup_{y \in Y} |\partial^\beta f(x, y)|.$$

Thus we can bound the integrals in the third term above, III, as follows,

$$\left| \int_{\Omega_\mu} K_\mu(x) R_\beta(x) x^\beta dx \right| \leq \sup_{x \in \Omega_\mu} |R_\beta(x) x^\beta| \|K\|_{L^1} \leq C \sup_{x \in \Omega_\mu} \sup_{y \in Y} |\partial^\beta f(x, y)| \mu^{|\beta|}.$$

Therefore,

$$(4.7) \qquad |III| \leq C \max_{|\beta| = p+1} \sup_{x \in \Omega_\mu} \sup_{y \in Y} |\partial^\beta f(x, y)| \mu^{p+1}.$$

Combining the estimates (4.5), (4.6) and (4.7) then yields the estimate in the lemma for functions with $\partial_x^\beta f \in C(\Omega, L^\infty(Y))$.

If we instead have that $\partial_x^\beta f \in L^\infty(\Omega_\mu; L^\infty(Y))$ for $0 \leq |\beta| \leq p + 1$, we can approximate them by smooth functions such that the above still holds. □

4.3. Averaging in space and slow time. For a vector-valued function $\mathbf{w}(x, t)$, let

$$(4.8) \qquad \bar{\mathcal{K}}_{\mu,\eta} \mathbf{w} := \int_{\Omega_\mu} \int_0^\eta K_\mu(x) K_\eta^0(t) \mathbf{w}(x, t) dt dx,$$

where $K \in \mathbb{K}^{p_x, q_x}$ and $K^0 \in \mathbb{K}_0^{p_t, q_t}$ are given kernels that are scaled by parameters μ and η, respectively.

LEMMA 4.5. *With $\bar{\mathcal{K}}_{\mu,\eta}$ given in (4.8), it holds for $\mathbf{u} \in L^\infty(0, \eta; L^2(\Omega))$ that*

$$(4.9) \qquad \left|\bar{\mathcal{K}}_{\mu,\eta} \mathbf{u}\right| \leq \frac{C}{\mu^{d/2}} \sup_{0 \leq t \leq \eta} \|\mathbf{u}(\cdot, t)\|_{L^2}.$$

Moreover, if $\mathbf{u} \in L^\infty(0, \eta; L^\infty(\Omega))$, then

$$(4.10) \qquad \left|\bar{\mathcal{K}}_{\mu,\eta} \mathbf{u}\right| \leq C \sup_{0 \leq t \leq \eta} \|\mathbf{u}(\cdot, t)\|_{L^\infty}.$$

Proof. By the Cauchy-Schwarz inequality, it follows that that

$$\left| \int_{\Omega_\mu} K_\mu(x) \mathbf{u}(x, t) dx \right| \leq \left(\int_{\Omega_\mu} \frac{1}{\mu^{2d}} \left| K\left(\frac{x}{\mu}\right) \right|^2 dx \int_{\Omega_\mu} |\mathbf{u}(x, t)|^2 dx \right)^{1/2}$$

$$= \left(\frac{1}{\mu^d} \int_{[-1,1]^d} |K(x)|^2 dx \int_{[-1,1]^d} |\mathbf{u}|^2 dx \right)^{1/2}$$

$$\leq \frac{C}{\mu^{d/2}} \|\mathbf{u}(\cdot, t)\|_{L^2},$$

hence it holds that

$$|\bar{\mathcal{K}}_{\mu,\eta}\mathbf{u}| \leq \frac{1}{\sqrt{\mu^d}}C \left\| \int_0^\eta K_\eta^0 \mathbf{u}(\cdot,t)dt \right\|_{L^2} \leq \frac{1}{\sqrt{\mu^d}}C\|K^0\|_{L^1} \sup_{0\leq t\leq \eta}\|\mathbf{u}(\cdot,t)\|_{L^2},$$

which shows the first result in the lemma. Furthermore, it holds that

$$|\bar{\mathcal{K}}_{\mu,\eta}\mathbf{u}| \leq \sup_{0\leq t\leq \eta}\|\mathbf{u}(\cdot,t)\|_{L^\infty(\Omega_\mu)} \int_{\Omega_\mu} \int_0^\eta |K_\mu||K_\eta^0|dtdx$$

$$\leq \|K\|_{L^1}\|K^0\|_{L^1} \sup_{0\leq t\leq \eta}\|\mathbf{u}(\cdot,t)\|_{L^\infty(\Omega_\mu)}.$$

This completes the proof. $\qquad\square$

Next, we prove a general lemma that holds for the averaging of sufficiently regular functions that change only slowly in time but contain both slow and fast variations in space. These fast spatial oscillations have to be representable by a periodic function that multiplies a function only depending on the slow variables.

LEMMA 4.6. *Consider averaging kernels $K(x)$ for space and $K^0(t)$ for time such that $K \in \mathbb{K}^{p_x,q_x}$ and $K^0 \in \mathbb{K}_0^{q_t,p_t}$ and assume that $\varepsilon < \mu < 1$ and $\varepsilon^2 < \eta < 1$. Let g be a 1-periodic function such that $g \in L^\infty(Y)$ and $\bar{g} = \int_{[0,1]^d} g(x)dx$. Suppose that for $0 \leq k \leq p_t + 1$ and $0 \leq |\beta| \leq p_x + 1$, $\partial_t^k \partial_x^\beta f(x,t) \in C(0,\eta; L^\infty(\Omega))$. Then*

$$\left|\bar{\mathcal{K}}_{\mu,\eta}\left(f(x,t)g(x/\varepsilon)\right) - f(0,0)\bar{g}\right|$$

$$\leq C \max_{k\leq p_t+1} \sup_{0\leq t\leq \eta} \|\partial_t^k f(\cdot,t)\|_{W^{p_x+1,\infty}(\Omega_\mu)}\|g\|_{L^\infty}\left(\left(\frac{\varepsilon}{\mu}\right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1}\right).$$

Proof. Consider first averaging in time only. As an immediate consequence of Definition 4.1 and the fact that $K^0(t)$ is zero for $t < 0$, it holds that

$$\int_0^\eta K_\eta^0(t)f(x,0)dt = f(x,0)\int_{-1}^1 K^0(t)dt = f(x,0),$$

$$\int_0^\eta K_\eta^0(t)\partial_t^j f(x,0)t^j dt = \eta^j \partial_t^j f(x,0)\int_{-1}^1 K^0(t)t^j dt = 0, \qquad 1 \leq j \leq p_t.$$

Hence, when Taylor-expanding $f(x,t)$ in time around zero, we obtain

$$\int_0^\eta K_\eta^0(t)f(x,t)dt = f(x,0) + \int_0^\eta K_\eta^0(t)R_{p_t+1}(x,t)dt,$$

where the remainder term is

$$R_{p_t+1}(x,t) := \frac{1}{p_t!}\int_0^t (t-z)^{p_t}\partial_t^{p_t+1}f(x,z)dz.$$

This representation of the time averaging integral can then be used to obtain a bound on the considered averaging error that consists of two parts,

$$\left|\bar{\mathcal{K}}_{\mu,\eta}\left(f(x,t)g(x/\varepsilon)\right) - f(0,0)\bar{g}\right|$$

$$\leq \left|\int_{\Omega_\mu} K_\mu(x)f(x,0)g(x/\varepsilon)dx - f(0,0)\bar{g}\right| + \left|\bar{\mathcal{K}}_{\mu,\eta}\left(R_{p_t}(x,t)g(x/\varepsilon)\right)\right| =: I + II.$$

The first part here, I, corresponds to averaging in space of a time-independent function with slow and fast, periodic variations in space. Application of Lemma 4.4 then yields

$$|I| \leq C \sup_{y \in Y} \|f(\cdot, 0)g(y)\|_{W^{p_x+1,\infty}(\Omega_\mu)} \left(\left(\frac{\varepsilon}{\mu} \right)^{q_x+2} + \mu^{p_x+1} \right).$$

To bound the second part, II, note that the remainder integral from time integration can be rescaled to $[0, 1]$ and then be bounded in terms of C_f and $K^0(t)$,

$$\int_0^\eta K_\eta^0 R(x,t)dt \leq \frac{1}{p_t!} \int_0^1 |K^0| \int_0^{\eta t} |(\eta t - z)|^{p_t} |\partial_t^{p_t+1} f(x,z)| dz dt$$

$$\leq \frac{\eta^{p_t+1}}{p_t!} \sup_{0 \leq t \leq \eta} |\partial_t^{p_t+1} f(x,t)| \|K^0\|_{L^1}.$$

Hence, we can bound the integral in II by

$$|II| \leq C \sup_{0 \leq t \leq \eta} \sup_{x \in \Omega_\mu} |\partial^{p_t+1} f(x,t)| \|g\|_{L^\infty} \eta^{p_t+1},$$

where the constant C depends on K, K^0 and p_t but is independent of ε, η and μ. Together with the estimate for $|I|$, this shows result in the lemma. $\square$

4.4. Averaging involving temporal oscillations. For expressions involving the correction term $\mathbf{v}\left(x, x/\varepsilon, t, t/\varepsilon^2\right)$, a special averaging lemma that exploits the structure of $\mathbf{v}$ as given in Lemma 3.2 is necessary in order to get error estimates in Section 5. We therefore proceed to derive a lemma for time averaging with a kernel for functions of the form

$$(4.11) \qquad \int_{-\mu}^{\mu} \partial_x^{\beta_1} f(x/\varepsilon) \partial_x^{\beta_2} K_\mu(x) u(x,t) \Psi \left(\frac{x}{\varepsilon}, \frac{t}{\varepsilon^2} \right) dx,$$

where $u(x,t)$ is a function that only varies slowly, $f(y)$ is a 1-periodic function and $\Psi(x,t)$ is defined as in Lemma 3.2. Moreover, β_1 and β_2 are given multi-indices. We obtain the following result.

LEMMA 4.7. *Consider averaging kernels $K \in \mathbb{K}^{q_x,p_x}$ and $K^0 \in \mathbb{K}_0^{q_t,p_t}$ and assume that $\varepsilon < \mu < 1$ and $\varepsilon^2 < \eta < 1$. Let $u(x,t)$ be a complex-valued function such that $\partial_t^r u \in C(0, \eta; L^\infty(\Omega))$ for $0 \leq r \leq q_t + 1$. Moreover, consider multi-indices β_1 and β_2 such that $|\beta_2| \leq q_x$. Suppose f is a 1-periodic function such that $\partial_x^{\beta_1} f \in L^\infty(Y)$, and that Ψ is given by (4.12). Then*

$$\left| \int_0^\eta K_\eta^0(t) \int_{\Omega_\mu} \left(\partial_x^{\beta_1} f(x/\varepsilon) \right) \left(\partial_x^{\beta_2} K_\mu(x) \right) u(x,t) \Psi \left(\frac{x}{\varepsilon}, \frac{t}{\varepsilon^2} \right) dx dt \right|$$

$$\leq C \frac{1}{\mu^{|\beta_2|} \varepsilon^{|\beta_1|}} \max_{\rho \leq q_t+1} \sup_{t \in [0,\eta]} \|\partial_t^\rho u(\cdot, t)\|_{L^\infty} \|\partial_y^{\beta_1} f\|_{L^\infty} \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1},$$

where the constant C depends on $\|K^0\|_{W^{1,q_t+1}}$ and $\|\boldsymbol{\chi}\|_{L^2}$ but is independent of ε, μ, η and t.

Proof. To simplify notation in the following, we introduce $\psi_j(t)$ such that

$$(4.12) \qquad \Psi(x,t) := \sum_{j=1}^\infty \boldsymbol{\chi}_j \phi_j(x) \psi_j(t), \qquad \text{where} \quad \psi_j(t) := e^{(-\alpha+i)\omega_j t}.$$

As in Lemma 3.2, ϕ_j and ω_j, $j \geq 0$ are the eigenfunctions and corresponding eigen-values of the operator $-L_{yy}$ and $\boldsymbol{\chi}_j$ are the expansion coefficients one obtains when expressing $\boldsymbol{\chi}$ in terms of the eigenfunction basis ϕ_j. Note that the time derivatives of ψ_j can be expressed in terms of the original function times a constant,

$$\frac{d}{dt}\psi_j\left(\frac{\eta t}{\varepsilon^2}\right) = -\frac{1}{c_j}\psi_j\left(\frac{\eta t}{\varepsilon^2}\right), \quad \text{with} \quad c_j := \frac{\varepsilon^2}{(\alpha - i)\omega_j \eta}.$$

Repeated application of integration by parts thus yields

$$\int_0^\eta K_\eta^0(t)u(x,t)\psi_j(t/\varepsilon^2)dt = -c_j\int_0^1 K^0(t)u(x,\eta t)\frac{d}{dt}\psi_j\left(\frac{\eta t}{\varepsilon^2}\right)dt$$

$$(4.13) \qquad\qquad\qquad = c_j^{q+1}\int_0^1 \partial_t^{q+1}\left(K^0(t)u(x,\eta t)\right)\psi_j\left(\frac{\eta t}{\varepsilon^2}\right)dt.$$

Since $\eta < 1$, we can moreover bound the absolute value of $\partial_t^{q+1}(K^0(t)u(x,\eta t))$ as

$$(4.14) \qquad\qquad \left|\partial_t^{q+1}(K^0(t)u(x,\eta t))\right| \leq C\max_{\rho \leq q_t+1}\sup_{t \in [0,\eta]}|\partial_t^\rho u(x,t)|.$$

Using the definition of Ψ, (4.12), the equality (4.13) and the fact that the averaging kernel $K^0(t)$ is zero for $t < 0$ since $K^0(t) \in \mathbb{K}_0^{q_t,p_t}$, we thus find that

$$(4.15) \quad \int_0^\eta K_\eta^0(t)u(x,t)\Psi\left(\frac{x}{\varepsilon},\frac{t}{\varepsilon^2}\right)dt = \sum_{j=1}^\infty \boldsymbol{\chi}_j\phi_j\left(\frac{x}{\varepsilon}\right)\int_0^\eta K_\eta^0(t)u(x,t)\psi_j\left(\frac{t}{\varepsilon^2}\right)dt$$

$$= \int_0^1 \partial_t^{q+1}\left(K^0(t)u(x,\eta t)\right)g^\varepsilon(x,\eta t)dt,$$

where

$$g(x,t) := \sum_{j=1}^\infty c_j^{q+1}\boldsymbol{\chi}_j\phi_j(x)\psi_j(t) \qquad \text{and} \qquad g^\varepsilon(x,t) = g(x/\varepsilon,t/\varepsilon^2).$$

Furthermore, we let for shortness of notation, $k_\mu^\varepsilon(x) := \left(\partial_x^{\beta_1}f(x/\varepsilon)\right)\left(\partial_x^{\beta_2}K_\mu(x)\right)$. It then follows by (4.14) and (4.15) that

$$\left|\int_0^\eta K_\eta^0(t)\int_{\Omega_\mu}\left(\partial_x^{\beta_1}f(x/\varepsilon)\right)\left(\partial_x^{\beta_2}K_\mu(x)\right)u(x,t)\Psi\left(x/\varepsilon,t/\varepsilon^2\right)dxdt\right|$$

$$= \left|\int_{\Omega_\mu}k_\mu^\varepsilon(x)\int_0^\eta K_\eta^0(t)u(x,t)\Psi\left(x/\varepsilon,t/\varepsilon^2\right)dtdx\right|$$

$$\leq \int_0^1\int_{\Omega_\mu}\left|\partial_t^{q+1}\left(K^0(t)u(x,\eta t)\right)\right|\left|k_\mu^\varepsilon(x)g^\varepsilon(x,\eta t)\right|dxdt$$

$$\leq C\max_{\rho \leq q_t+1}\sup_{t \in [0,\eta]}\|\partial_t^\rho u(\cdot,t)\|_{L^\infty}\int_0^1\int_{\Omega_\mu}\left|k_\mu^\varepsilon(x)g^\varepsilon(x,\eta t)\right|dxdt.$$

Rescaling of the spatial integral and application of the Cauchy-Schwarz inequality

yields

$$\int_{\Omega_\mu} |k_\mu^\varepsilon(x) g^\varepsilon(x, \eta t))| dx = \int_{[-1,1]^d} \left| \mu^{-|\beta_1|-|\beta_2|} \partial_x^{\beta_1} f(\mu x/\varepsilon) \partial_x^{\beta_2} K(x) g^\varepsilon(\mu x, \eta t) \right| dx$$

$$\leq \left(\int_{[-1,1]^d} |\mu^{-|\beta_1|-|\beta_2|} \partial_x^{\beta_1} f(\mu x/\varepsilon) \partial_x^{\beta_2} K(x)|^2 dx \right)^{1/2} \left(\int_{[-1,1]^d} |g^\varepsilon(\mu x, \eta t)|^2 dx \right)^{1/2}.$$

Note that $g(x, t)$ is 1-periodic in space and $\mu > \varepsilon$, which implies that the latter integral can be bounded by the corresponding $\| \cdot \|_{L^2(Y)}$-norm using Lemma 4.8 below,

$$\int_{[-1,1]^d} |g^\varepsilon(\mu x, \eta t)|^2 dx = \int_{[-1,1]^d} \left| g\left(\mu x/\varepsilon, \eta t/\varepsilon^2 \right) \right|^2 dx \leq C \left\| g\left(\cdot, \eta t/\varepsilon^2 \right) \right\|_{L^2(Y)}^2.$$

Since the absolute values of the eigenvalues ω_j are increasing with j, all c_j, $j \geq 1$ can be bounded by c_1, the constant involving ω_1 the smallest non-zero eigenvalue

$$(4.16) \quad |c_j| = \left| \frac{\varepsilon^2}{(\alpha - i)\omega_j \eta} \right| = \frac{1}{\omega_j} \frac{\varepsilon^2}{\eta \sqrt{1 + \alpha^2}} \leq \frac{1}{\omega_1} \frac{\varepsilon^2}{\eta \sqrt{1 + \alpha^2}} = |c_1|, \qquad j = 1, 2, \dots.$$

One can therefore show using the orthogonality of the basis functions ϕ_j and the boundedness of ψ_j, that

$$\left\| g\left(\cdot, \eta t/\varepsilon^2 \right) \right\|_{L^2(Y)}^2 = \sum_{j=1}^\infty \left| c_j^{q_t+1} \chi_j \psi_j \left(\eta t/\varepsilon^2 \right) \right|^2 \leq |c_1|^{2(q_t+1)} \sum_{j=1}^\infty |\chi_j|^2$$

$$= |c_1|^{2(q_t+1)} \|\chi\|_{L^2(Y)}^2,$$

for any time $t \geq 0$. It furthermore holds that

$$\int_{[-1,1]^d} |\mu^{-|\beta_1|-|\beta_2|} \partial_x^{\beta_1} f(\mu x/\varepsilon) \partial_x^{\beta_2} K(x)|^2 dx \leq \frac{1}{\mu^{2|\beta_2|} \varepsilon^{2|\beta_1|}} \sup_{y \in Y} |\partial_y^{\beta_1} f(y)|^2 \|K\|_{H^{|\beta_2|}}^2.$$

Therefore, we obtain

$$\int_0^1 \int_{\Omega_\mu} |k_\mu^\varepsilon(x) g^\varepsilon(x, \eta t))| dx dt \leq \|\chi\|_{L^2} \|K\|_{H^{|\beta_2|}} \|K^0\|_{W^{1, q_t+1}} \frac{|c_1|^{(q_t+1)}}{\mu^{|\beta_2|} \varepsilon^{|\beta_1|}} \sup_{y \in Y} |\partial_y^{\beta_1} f(y)|$$

and the result in the lemma follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

In order to derive the above result, we used the following lemma from [10], which is a useful tool when working with periodic functions. For the convenience of the reader, a short proof is given here.

LEMMA 4.8. *Assume $g \in L^2(Y)$ is 1-periodic. Let $\theta \geq 1$ and $a < b$ be given real constants, then*

$$\int_{[a,b]^d} |g(\theta x)|^2 dx \leq C \|g\|_{L^2(Y)}^2,$$

where C only depends on a, b and the dimension d but is independent of θ.

Proof. Let $K = \lfloor (b-a)\theta \rfloor$, the number of full periods of $g(\theta x)$ in the interval $[a, b]$ (in one coordinate direction). Consider first a rescaled integral in the jth coordinate direction. As $a\theta + (K + 1) > b\theta$ and $|g|^2 > 0$, it holds that

$$\int_{a\theta}^{b\theta} |g(x)|^2 dx_j \leq \sum_{k=0}^K \int_{a\theta+k}^{a\theta+(k+1)} |g(x)|^2 dx_j = (K + 1) \int_0^1 |g(x)|^2 dx_j.$$

Therefore, we find that

$$\int_{[a,b]^d} |g(\theta x)|^2 dx = \frac{1}{\theta^d} \int_{a\theta}^{b\theta} \cdots \int_{a\theta}^{b\theta} |g(x)|^2 dx_1 \cdots dx_d$$

$$\leq \frac{K+1}{\theta^d} \int_{a\theta}^{b\theta} \cdots \int_0^1 |g(x)|^2 dx_1 \cdots dx_d \leq \left(\frac{K+1}{\theta}\right)^d \|g\|_{L^2(Y)}^2.$$

Since $\theta \geq 1$,

$$\frac{K+1}{\theta} \leq b - a + \frac{1}{\theta} \leq b - a + 1,$$

which entails that the constant multiplying $\|g\|_{L^2(Y)}^2$ is independent of θ. $\qquad \square$

5. HMM approximation errors. In this section, we prove bounds for the averaging error in each of the three models (M1), (M2) and (M3) described in Section 2. These error bounds depend on the parameters of the kernels used for averaging in time and space, p_x, q_x, p_t and q_t as well as the sizes of the averaging domains. Given a sufficiently regular solution $\mathbf{m}_0$, which we assume in (A5), choosing high values for the kernel parameters makes it possible to reduce the averaging error to $\mathcal{O}(\varepsilon)$ as stated in the following theorem.

THEOREM 5.1. *Assume (A1)-(A5) hold and let $\mathbf{A}^H$ be the homogenized coefficient matrix corresponding to $a^\varepsilon(x)$, given in (1.6). Consider $\bar{\mathcal{K}}_{\eta,\mu}$ with averaging kernels $K \in \mathbb{K}^{p_x,q_x}$ and $K^0 \in \mathbb{K}_0^{p_t,q_t}$ and let $\varepsilon < \mu < 1$ and $\varepsilon^2 < \eta < \min(1, T^\varepsilon)$, where T^ε is given by (3.2). Then*

$$\left|\bar{\mathcal{K}}_{\mu,\eta}(a^\varepsilon \boldsymbol{\nabla}\mathbf{m}^\varepsilon) - \boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H\right|$$
$$\leq C\left(\varepsilon + \left(\frac{\varepsilon}{\mu}\right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1} + \left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}\right),$$

$$\left|\bar{\mathcal{K}}_{\mu,\eta}(\boldsymbol{\nabla}\cdot(a^\varepsilon \boldsymbol{\nabla}\mathbf{m}^\varepsilon)) - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)\right|$$
$$\leq C\left(\varepsilon + \left(\frac{\varepsilon}{\mu}\right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1} + \frac{1}{\mu}\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}\right),$$

and

$$\left|\bar{\mathcal{K}}_{\mu,\eta}(\mathbf{m}^\varepsilon \times \boldsymbol{\nabla}\cdot(a^\varepsilon \boldsymbol{\nabla}\mathbf{m}^\varepsilon)) - \mathbf{m}_0(0,0) \times \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)\right|$$
$$\leq C\left(\varepsilon + \left(\frac{\varepsilon}{\mu}\right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1} + \frac{1}{\mu}\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}\right).$$

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

Given that the initial data to the micro problem is chosen such that $\partial_x^\beta \mathbf{m}(0,0)$, $|\beta| \leq 2$ agree with the corresponding derivatives of the macro solution $\mathbf{M}$ at the point in time and space that one averages around, Theorem 5.1 provides estimates for E_i, $i = 1, 2, 3$ as given in Section 2.

To prove the estimates in Theorem 5.1, we consider an approximation $\mathbf{m}^{\text{app}} := \mathbf{m}_0(x,t) + \varepsilon\mathbf{m}_1(x, x/\varepsilon, t, t/\varepsilon^2)$ to $\mathbf{m}^\varepsilon$. We then proceed in a similar way for all three models. We first show that averaging of $\mathbf{m}^{\text{app}}$ results in approximations to the quantities required to complete the models up to a certain error. The contribution of

$\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}}$ only gives a remainder term resulting in an additional error. More precisely, it holds for the approximation error in the first model, (M1), that

$$
(5.1) \quad
\begin{aligned}
|\bar{\mathcal{K}}_{\mu,\eta}(a^\varepsilon \boldsymbol{\nabla}\mathbf{m}^\varepsilon) - \boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H| &\leq |\bar{\mathcal{K}}_{\mu,\eta}(a^\varepsilon \boldsymbol{\nabla}\mathbf{m}^{\mathrm{app}}) - \boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H| \\
&\quad + |\bar{\mathcal{K}}_{\mu,\eta}(a^\varepsilon \boldsymbol{\nabla}(\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}})| \\
&=: e_{M1} + r_{M1}.
\end{aligned}
$$

Similarly, we have for (M2) that

$$
(5.2) \quad
\begin{aligned}
&|\bar{\mathcal{K}}_{\mu,\eta}(\mathcal{L}\mathbf{m}^\varepsilon) - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)| \\
&\qquad \leq |\bar{\mathcal{K}}_{\mu,\eta}(\mathcal{L}\mathbf{m}^{\mathrm{app}}) - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)| + |\bar{\mathcal{K}}_{\mu,\eta}(\mathcal{L}(\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}}))| \\
&\qquad =: e_{M2} + r_{M2}\,,
\end{aligned}
$$

and in case of (M3)

$$
(5.3) \quad
\begin{aligned}
&|\bar{\mathcal{K}}_{\mu,\eta}\left(\mathbf{m}^\varepsilon \times \boldsymbol{\nabla}\cdot(a^\varepsilon\boldsymbol{\nabla}\mathbf{m}^\varepsilon)\right) - \mathbf{m}_0(0,0) \times \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)| \\
&\qquad \leq |\mathbf{m}^{\mathrm{app}} \times \mathcal{L}\mathbf{m}^{\mathrm{app}} - \mathbf{m}_0(0,0) \times \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)| \\
&\qquad\quad + |\mathbf{m}^{\mathrm{app}} \times \mathcal{L}(\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}}) + (\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}}) \times \mathcal{L}\mathbf{m}^{\mathrm{app}}| =: e_{M3} + r_{M3}\,.
\end{aligned}
$$

Each of the approximation errors e_{Mi}, $i = 1, 2, 3$, is then bounded using Lemma 4.4 – Lemma 4.7 as shown in the following sections. Finally, estimates for the norms of the remainder terms r_{Mi} are given in Section 5.4, which completes the proof of Theorem 5.1.

For the derivations, we define the y-periodic functions

$$
(5.4) \qquad \mathbf{g}(y) := a(y)(\mathbf{I} + \boldsymbol{\nabla}_y\boldsymbol{\chi}(y))\,, \quad \text{and} \quad \mathbf{h}(y) = a(y)\boldsymbol{\chi}(y)\,.
$$

Since by assumption $(A1)$, $a \in C^\infty(Y)$, the same holds for $\boldsymbol{\chi}$ and thus also $\mathbf{g}, \mathbf{h} \in C^\infty(Y)$. Note that $\mathbf{g}$ is a matrix-valued function, $\mathbf{g} : \mathbb{R} \to \mathbb{R}^{d \times d}$, and we denote its elements by $\mathbf{g}_{ij}$. By the definition of the homogenized matrix $\mathbf{A}^H$ in (1.6), it holds that

$$
(5.5) \qquad \mathbf{A}_{ij}^H = \int_Y \mathbf{g}_{ij}(y)dy\,.
$$

The average of $\mathbf{h}$ is in general non-zero. In the following, we use the notation $\mathbf{g}^\varepsilon := \mathbf{g}(x/\varepsilon), \mathbf{h}^\varepsilon := \mathbf{h}(x/\varepsilon)$ and $\boldsymbol{\chi}^\varepsilon := \boldsymbol{\chi}(x/\varepsilon)$. Moreover, we define

$$
\mathbf{u}(x, y, t) := \mathbf{m}_0(x, t) + \varepsilon\boldsymbol{\nabla}_x\mathbf{m}_0(x, t)\boldsymbol{\chi}(y),
$$

and let $\mathbf{u}^\varepsilon := \mathbf{u}(x, x/\varepsilon, t)$. Together with (3.6), this implies that

$$
(5.6) \qquad \mathbf{m}^{\mathrm{app}} = \mathbf{m}_0(x, t) + \varepsilon\mathbf{m}_1(x, x/\varepsilon, t, t/\varepsilon^2) = \mathbf{u}^\varepsilon + \varepsilon\mathbf{v}^\varepsilon,
$$

where $\mathbf{v}^\varepsilon = \mathbf{v}(x, x/\varepsilon, t, t/\varepsilon^2)$, for $\mathbf{v}$ as given by Lemma 3.2, and

$$
(5.7) \qquad a^\varepsilon\boldsymbol{\nabla}\mathbf{u}(x, x/\varepsilon, t) = \boldsymbol{\nabla}_x\mathbf{m}_0(x, t)\mathbf{g}(x/\varepsilon) + \varepsilon\boldsymbol{\nabla}_x(\boldsymbol{\nabla}_x\mathbf{m}_0(x, t)\mathbf{h}(x/\varepsilon)).
$$

Furthermore, it follows from the definition of $\boldsymbol{\chi}$ via the cell problem, that the divergence of $\mathbf{g}$ is zero,

$$
(5.8) \quad \boldsymbol{\nabla}_y \cdot \mathbf{g}(y) = \boldsymbol{\nabla}_y \cdot (a(y)(\mathbf{I} + \boldsymbol{\nabla}_y\boldsymbol{\chi}(y))) = \boldsymbol{\nabla}_y a(y) + \boldsymbol{\nabla}_y \cdot (a(y)\boldsymbol{\nabla}_y\boldsymbol{\chi}(y)) = \mathbf{0}\,.
$$

Hence, we have

$$\mathcal{L}\mathbf{m}^{\mathrm{app}} = \boldsymbol{\nabla} \cdot \left(a(x/\varepsilon)\boldsymbol{\nabla}\left(\mathbf{u}(x,x/\varepsilon,t) + \varepsilon\mathbf{v}(x,x/\varepsilon,t,t/\varepsilon^2)\right)\right)$$

$$(5.9) \qquad = \boldsymbol{\nabla}_x \cdot (\boldsymbol{\nabla}_x\mathbf{m}_0\mathbf{g}^\varepsilon) + \boldsymbol{\nabla}_y \cdot (\boldsymbol{\nabla}_x(\boldsymbol{\nabla}_x\mathbf{m}_0\mathbf{h}^\varepsilon)) + \varepsilon\boldsymbol{\nabla}_x \cdot (\boldsymbol{\nabla}_x(\boldsymbol{\nabla}_x\mathbf{m}_0\mathbf{h}^\varepsilon)) + \varepsilon\mathcal{L}\mathbf{v}^\varepsilon.$$

Finally, for matters of brevity, we define a short-hand notation for the error terms in Lemma 4.4 and Lemma 4.6,

$$(5.10) \qquad\qquad e_{44}(\varepsilon,\mu) := \left(\frac{\varepsilon}{\mu}\right)^{q_x+2} + \mu^{p_x+1},$$

$$(5.11) \qquad\qquad e_{46}(\varepsilon,\mu,\eta) := \left(\frac{\varepsilon}{\mu}\right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1}.$$

5.1. Approximating $\boldsymbol{\nabla}\mathbf{m}_0\mathbf{A}^H$. We now apply the lemmas from the previous sections to obtain an estimate for the error e_{M1} as given by (5.1).

LEMMA 5.2. *Under the assumptions in Theorem 5.1, it holds that*

$$|\bar{\mathcal{K}}_{\mu,\eta}(a^\varepsilon\boldsymbol{\nabla}\mathbf{m}^{\mathrm{app}}) - \boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H|$$

$$\leq C\left(\left(\frac{\varepsilon}{\mu}\right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1} + \left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1} + \varepsilon\right),$$

where C is independent of ε,μ and η.

Proof. We start by splitting e_{M1} according to (5.6), into a part without fast oscillations in time and a part containing $\mathbf{v}$,

$$(5.12) \qquad e_{M1} \leq \left|\bar{\mathcal{K}}_{\mu,\eta}(a^\varepsilon\boldsymbol{\nabla}\mathbf{u}^\varepsilon) - \boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H\right| + \left|\bar{\mathcal{K}}_{\mu,\eta}(\varepsilon a^\varepsilon\boldsymbol{\nabla}\mathbf{v}^\varepsilon)\right| = I_1 + I_2.$$

In the following we use the notation $\mathbf{g}_\beta$ and $\mathbf{A}_\beta^H$ to refer to the element in the corresponding matrix that multiplies $\partial^\beta\mathbf{m}_0$ according to (5.7) and similarly for $\mathbf{h}$. It then follows by (5.7) that there are constants c_β, d_β such that

$$a^\varepsilon\boldsymbol{\nabla}\mathbf{u}^\varepsilon = \sum_{|\beta|=1} c_\beta\partial_x^\beta\mathbf{m}_0(x,t)\mathbf{g}_\beta(x/\varepsilon) + \varepsilon\sum_{|\beta|=2} d_\beta\partial_x^\beta\mathbf{m}_0(x,t)\mathbf{h}_\beta(x/\varepsilon).$$

Since by assumption (A5), we have $\mathbf{m}_0(x,t) \in C^\infty(0,T;H^\infty(\Omega))$, an application of Lemma 4.6 together with (5.5) yields that given a multi-index β,

$$(5.13) \qquad\qquad |\bar{\mathcal{K}}_{\mu,\eta}(\partial_x^\beta\mathbf{m}_0\mathbf{g}_\beta^\varepsilon) - \partial_x^\beta\mathbf{m}_0(0,0)\mathbf{A}_\beta^H| \leq Ce_{46}(\varepsilon,\mu,\eta).$$

As furthermore by the boundedness of $\bar{\mathcal{K}}_{\eta,\mu}$, (4.10) in Lemma 4.5,

$$|\bar{\mathcal{K}}_{\mu,\eta}(\partial_x^\beta\mathbf{m}_0\mathbf{h}_\beta^\varepsilon)| \leq C,$$

(5.13) implies that

$$I_1 \leq C\left(\varepsilon + e_{46}(\varepsilon,\mu,\eta)\right).$$

The remaining term in (5.12), I_2, is rewritten using integration by parts, which together with the definition of $\mathbf{v}$ according to Lemma 3.2 yields

$$I_2 = \varepsilon\left|\mathcal{R}e\left(\int_{\Omega_\mu}\int_{-\eta}^\eta K_\eta^0(t)\nabla\left(a\left(x/\varepsilon\right)K_\mu(x)\right)\mathbf{f}_\mathbf{v}(x,t)\Psi\left(x/\varepsilon,t/\varepsilon^2\right)dtdx\right)\right|.$$

The regularity assumption (A5) for $\mathbf{m}_0$ implies that $\mathbf{f_v} \in C^\infty(0, T; H^\infty(\Omega))$ and

$$\max_{r \leq q_t+1} \sup_{t \in [0,\eta]} \|\partial_t^r \mathbf{f}_v(\cdot, t)\|_{L^\infty} \leq C,$$

wherefore it follows by Lemma 4.7 that

$$|I_2| \leq C \left(\frac{\varepsilon}{\mu} \|a\|_{L^\infty} + \|\partial_y a\|_{L^\infty} \right) \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1} \leq C \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1}.$$

Hence, we obtain the estimate in the lemma. $\qquad\square$

5.2. Approximating $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_0 \mathbf{A}^H)$. To estimate the first part of the approximation error in the second model, e_{M2} as given in (5.2) we proceed in a similar way as before. However, as we consider the divergence of the gradient, more terms need to be estimated. This results in the following lemma.

LEMMA 5.3. *Under the assumptions in Theorem* 5.1 *it holds that*

$$|\bar{\mathcal{K}}_{\mu,\eta}(\boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla}\mathbf{m}^{\mathrm{app}})) - \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)|$$

$$\leq C \left(\left(\frac{\varepsilon}{\mu} \right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1} + \frac{1}{\mu} \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1} + \varepsilon \right),$$

where C is independent of ε, μ and η.

Proof. To begin with, we again split the error under consideration into two parts,

$$e_{M2} \leq \left| \bar{\mathcal{K}}_{\mu,\eta}(\mathcal{L}\mathbf{u}^\varepsilon) - \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H) \right| + |\bar{\mathcal{K}}_{\mu,\eta}(\varepsilon\mathcal{L}\mathbf{v}^\varepsilon)| =: II_1 + II_2.$$

Based on (5.9) one can deduce that there are constant coefficients c_β, d_β such that

$$(5.14) \quad \begin{aligned} \mathcal{L}\mathbf{u}^\varepsilon &= \mathcal{L}(\mathbf{m}_0 + \varepsilon\boldsymbol{\nabla}\mathbf{m}_0\boldsymbol{\chi}^\varepsilon) \\ &= \boldsymbol{\nabla}_x \cdot (\boldsymbol{\nabla}_x\mathbf{m}_0\mathbf{g}^\varepsilon) + \boldsymbol{\nabla}_y \cdot (\boldsymbol{\nabla}_x(\boldsymbol{\nabla}_x\mathbf{m}_0\mathbf{h}^\varepsilon)) + \varepsilon\boldsymbol{\nabla}_x \cdot (\boldsymbol{\nabla}_x(\boldsymbol{\nabla}_x\mathbf{m}_0\mathbf{h}^\varepsilon)) \\ &= \sum_{|\beta|=2} c_\beta[\mathbf{g}_\beta + (\boldsymbol{\nabla}_y\mathbf{h})_\beta](x/\varepsilon)\partial_x^\beta\mathbf{m}_0(x,t) + \varepsilon \sum_{|\beta|=3} d_\beta\mathbf{h}_\beta(x/\varepsilon)\partial_x^\beta\mathbf{m}_0(x,t). \end{aligned}$$

Similar to before, $\mathbf{g}_\beta$ here denotes the component of $\mathbf{g}$ that multiplies $\partial_x^\beta\mathbf{m}$ and accordingly for the other quantities. Hence, it holds that

$$\begin{aligned} II_1 \leq \sum_{|\beta|=2} c_\beta \left| \bar{\mathcal{K}}_{\mu,\eta}\left(\mathbf{g}_\beta^\varepsilon\partial_x^\beta\mathbf{m}_0\right) - \mathbf{A}_\beta^H\partial_x^\beta\mathbf{m}_0(0,0) \right| &+ \sum_{|\beta|=2} c_\beta \left| \bar{\mathcal{K}}_{\mu,\eta}\left((\boldsymbol{\nabla}_y\mathbf{h}^\varepsilon)_\beta\partial_x^\beta\mathbf{m}_0\right) \right| \\ &+ \varepsilon \sum_{|\beta|=3} d_\beta \left| \bar{\mathcal{K}}_{\mu,\eta}\left(\mathbf{h}_\beta^\varepsilon\partial^\beta\mathbf{m}_0\right) \right|. \end{aligned}$$

Note that since $\mathbf{h}$ is y-periodic, the average with respect to y of the second term on the right-hand side here is zero. The averages in the other two sums can be bounded using (5.13) and (4.10) in Lemma 4.5 and it follows in the same way as in the estimate of e_{M1} that

$$II_1 \leq_C (\varepsilon + e_{46}(\varepsilon, \mu, \eta)).$$

Furthermore, using integration by parts and Lemma 3.2, we can rewrite II_2 as

$$\begin{aligned} II_2 &= \varepsilon \left| \int_{-\eta}^\eta K_\eta^0(t) \int_{\Omega_\mu} (LK_\mu(x)) \mathbf{v}^\varepsilon dx dt \right| \\ &= \varepsilon \left| \mathcal{R}e \left(\int_{-\eta}^\eta K_\eta^0(t) \int_{\Omega_\mu} \nabla \cdot (a(x/\varepsilon)\nabla K_\mu) \mathbf{f_v}(x,t)\Psi \left(\frac{x}{\varepsilon}, \frac{t}{\varepsilon^2} \right) dx dt \right) \right|, \end{aligned}$$

Thus it follows by Lemma 4.7 with $|\beta_1| \leq 1$ and $|\beta_2| = 2 - |\beta_1|$ that

$$II_2 \leq C \left(\frac{\varepsilon}{\mu^2} \|a\|_\infty + \frac{1}{\mu} \|\partial_y a\|_\infty \right) \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1} \leq C \frac{1}{\mu} \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1},$$

which results in the estimate in Lemma 5.3.   □

5.3. Approximating $\mathbf{m}_0 \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_0 \mathbf{A}^H)$. We now consider e_{M3}, the first contribution to the error bound for the approximation error in the third model as given in (5.3). Since we are now considering a nonlinear expression, the derivations are more involved than for the previous two models. However, the resulting estimate is very similar to the one in Lemma 5.3, as stated in the following.

LEMMA 5.4. *Under the assumptions in Theorem 5.1 it holds that*

$$|\bar{\mathcal{K}}_{\mu,\eta}(\mathbf{m}^{\mathrm{app}} \times \mathcal{L}\mathbf{m}^{\mathrm{app}})) - \mathbf{m}_0(0,0) \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)|$$
$$\leq C \left(\left(\frac{\varepsilon}{\mu} \right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1} + \frac{1}{\mu} \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1} + \varepsilon \right),$$

where the constant C is independent of ε, μ and η.

Proof. For the term under consideration here, (5.6) implies that the error e_{M3} can be split into four parts,

$$\begin{aligned}
e_{M3} \leq &|\bar{\mathcal{K}}_{\mu,\eta}(\mathbf{u}^\varepsilon \times \mathcal{L}\mathbf{u}^\varepsilon) - \mathbf{m}_0(0,0) \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)| \\
&+ \varepsilon|\bar{\mathcal{K}}_{\mu,\eta}(\mathbf{u}^\varepsilon \times \mathcal{L}\mathbf{v}^\varepsilon)| + \varepsilon|\bar{\mathcal{K}}_{\mu,\eta}(\mathbf{v}^\varepsilon \times \mathcal{L}\mathbf{u}^\varepsilon)| + \varepsilon^2|\bar{\mathcal{K}}_{\mu,\eta}(\mathbf{v}^\varepsilon \times \mathcal{L}\mathbf{v}^\varepsilon)| \\
=: &III_1 + \varepsilon III_2 + \varepsilon III_3 + \varepsilon^2 III_4.
\end{aligned} \tag{5.15}$$

Using the sums in (5.14) to express $\mathcal{L}\mathbf{u}^\varepsilon$, we find

$$\tag{5.16}$$
$$\begin{aligned}
\mathbf{u}^\varepsilon \times \mathcal{L}\mathbf{u}^\varepsilon &= (\mathbf{m}_0 + \varepsilon \sum_{|\gamma|=1} \boldsymbol{\chi}_\gamma^\varepsilon \partial_x^\gamma \mathbf{m}_0) \times \mathcal{L}\mathbf{u}^\varepsilon \\
&= \sum_{|\gamma|\leq 1} \left(\sum_{|\beta|=2} c_\beta [(\varepsilon\boldsymbol{\chi}_\gamma)^{|\gamma|}(\mathbf{g} + \boldsymbol{\nabla}_y\mathbf{h})_\beta] (x/\varepsilon) [\partial_x^\gamma \mathbf{m}_0 \times \partial_x^\beta \mathbf{m}_0](x,t) \right. \\
&\qquad\qquad \left. + \varepsilon \sum_{|\beta|=3} d_\beta [(\varepsilon\boldsymbol{\chi}_\gamma)^{|\gamma|}\mathbf{h}_\beta] (x/\varepsilon) [\partial_x^\gamma \mathbf{m}_0 \times \partial_x^\beta \mathbf{m}_0](x,t) \right) \\
&=: \sum_{|\beta|=2} c_\beta (\mathbf{g} + \boldsymbol{\nabla}_y\mathbf{h})_\beta (x/\varepsilon) [\mathbf{m}_0 \times \partial_x^\beta \mathbf{m}_0](x,t) + \varepsilon\mathbf{f}_0^\varepsilon(x,t) + \varepsilon^2\mathbf{f}_1^\varepsilon(x,t),
\end{aligned}$$

for some functions $\mathbf{f}_0^\varepsilon, \mathbf{f}_1^\varepsilon \in L^\infty(0,\eta; L^\infty(\Omega))$. Since $\mathbf{m}_0 \in C^\infty(0,T; H^\infty(\Omega))$, we also have $\partial_x^\gamma \mathbf{m}_0 \times \partial_x^\beta \mathbf{m}_0 \in C^\infty(0,T; H^\infty(\Omega))$, and obtain using Lemma 4.6 that

$$|\bar{\mathcal{K}}_{\mu,\eta}([\mathbf{m}_0 \times \partial_x^\beta \mathbf{m}_0](x,t)\mathbf{g}_\beta(x/\varepsilon)) - [\mathbf{m}_0 \times \partial_x^\beta \mathbf{m}_0](0,0)\mathbf{A}_\beta^H| \leq C e_{46}(\varepsilon,\mu,\eta).$$

Furthermore, as $\mathbf{h}$ is y-periodic, the average of $\partial_y^\nu \mathbf{h}_\beta$ for $|\nu| = 1$ is zero, hence

$$\left| \bar{\mathcal{K}}_{\mu,\eta}([\mathbf{m}_0 \times \partial_x^\beta \mathbf{m}_0](x,t)\partial_y^\nu \mathbf{h}_\beta](x/\varepsilon)) \right| \leq C e_{46}(\varepsilon,\mu,\eta).$$

The remaining terms involved in (5.16) can be bounded using (4.10) in Lemma 4.5,

$$|\bar{\mathcal{K}}_{\mu,\eta}\mathbf{f}_0^\varepsilon| \le C, \qquad |\bar{\mathcal{K}}_{\mu,\eta}\mathbf{f}_1^\varepsilon| \le C.$$

In total, we can thus estimate III_1 as

$$(5.17) \qquad III_1 = \left|\bar{\mathcal{K}}_{\mu,\eta}(\mathbf{u}^\varepsilon \times \mathcal{L}\mathbf{u}^\varepsilon) - \mathbf{m}_0(0,0) \times \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\mathbf{m}_0(0,0)\mathbf{A}^H)\right|$$
$$\le C\left(\varepsilon + e_{46}(\varepsilon,\mu,\eta)\right).$$

The next two terms in (5.15) can be bounded using Lemma 4.7. Consider first εIII_2. Using integration by parts as in the previous sections to obtain $\mathbf{v}$ without any spatial derivatives yields

$$\int_{\Omega_\mu} K_\mu \mathbf{u}^\varepsilon \times \mathcal{L}\mathbf{v}^\varepsilon dx = \int_{\Omega_\mu} [\boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla}(K_\mu \mathbf{m}_0)) + \varepsilon \boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla}(K_\mu \boldsymbol{\nabla}\mathbf{m}_0 \boldsymbol{\chi}^\varepsilon))] \times \mathbf{v}^\varepsilon dx.$$

Let now $\beta = \beta_1 + \beta_2 + \beta_3$ a multi-index. Using Lemma 3.2, we can rewrite

$$\boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla}(K_\mu \mathbf{m}_0)) \times \mathbf{v}^\varepsilon$$
$$= -\mathcal{R}e\left(\sum_{|\beta_1|\le 1, |\beta|=2} c_\beta \partial_x^{\beta_1} a(x/\varepsilon)\partial_x^{\beta_2} K_\mu(x)[\partial_x^{\beta_3}\mathbf{m}_0 \times \mathbf{f}_v](x,t)\Psi(x/\varepsilon, t/\varepsilon^2)\right)$$

for some coefficients c_β that might also be zero. It therefore follows by Lemma 4.7, the regularity assumption (A5) and the fact that $\varepsilon < \mu$, that

$$\varepsilon \int_0^\eta K_\eta \int_{\Omega_\mu} \boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla}(K_\mu \mathbf{m}_0)) \times \mathbf{v}^\varepsilon dx dt$$
$$\le C \sum_{\substack{|\beta_1|\le 1, \\ |\beta_2|\le 2-|\beta_1|}} \frac{\varepsilon}{\mu^{|\beta_2|}\varepsilon^{|\beta_1|}}\|\partial_y^{\beta_1}a\|_{L^\infty}\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1} \le C\frac{1}{\mu}\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}.$$

Similarly, it holds with $\gamma = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$ that that

$$\boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla}(K_\mu \boldsymbol{\nabla}\mathbf{m}_0 \boldsymbol{\chi}^\varepsilon)) \times \mathbf{v}^\varepsilon$$
$$= -\mathcal{R}e\left(\sum_{\substack{|\gamma_1|\le 1 \\ |\gamma|\le 2}}\sum_{|\nu|=1} c_\gamma[\partial_x^{\gamma_1}a\partial_x^{\gamma_4}\boldsymbol{\chi}_\gamma](x/\varepsilon)\partial_x^{\gamma_2}K_\mu(x)[\partial_x^{\gamma_3}\partial_x^\nu \mathbf{m}_0 \times \mathbf{f}_v](x,t)\Psi(x/\varepsilon, t/\varepsilon^2)\right)$$

Application of Lemma 4.7, with $|\beta_1| = 0$ and $f(y) = \varepsilon^{-|\gamma_1|-|\gamma_4|}[\partial_y^{\gamma_1}a\partial_y^{\gamma_4}\boldsymbol{\chi}_\gamma](y)$, then yields

$$\varepsilon^2 \int_0^\eta K_\eta \int_{\Omega_\mu} \boldsymbol{\nabla} \cdot (a^\varepsilon \boldsymbol{\nabla}(K_\mu \mathbf{m}_0)) \times \mathbf{v}^\varepsilon dx dt$$
$$\le C \sum_{\substack{|\gamma_1|\le 1, \\ |\gamma_2|+|\gamma_4|\le 2-|\gamma_1|}} \left(\frac{\varepsilon^2}{\varepsilon^{|\gamma_1|+|\gamma_4|}\mu^{|\gamma_2|}}\|\partial_y^{\gamma_1}a\partial_y^{\gamma_4}\boldsymbol{\chi}\|_{L^\infty}\right)\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1} \le C\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}.$$

By (A1) and since $\varepsilon/\mu < 1$, we thus have in total

$$(5.18) \qquad \varepsilon III_2 = \left|\bar{\mathcal{K}}_{\mu,\eta}(\varepsilon\mathbf{u}^\varepsilon \times \mathcal{L}\mathbf{v}^\varepsilon)\right| \le C\frac{1}{\mu}\left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}.$$

The next term in (5.15), εIII_3, can be treated similarly. Expressing $\mathcal{L}\mathbf{u}^\varepsilon$ as given in (5.14) and using Lemma 3.2, we get

$$\mathcal{L}\mathbf{u}^\varepsilon \times \mathbf{v}^\varepsilon = -\mathcal{R}e\Big(\sum_{|\beta|=2} c_\beta [\mathbf{g}_\beta + (\boldsymbol{\nabla}_y \mathbf{h})_\beta](x/\varepsilon)[\partial_x^\beta \mathbf{m}_0 \times \mathbf{f}_v](x,t)\Psi\left(x/\varepsilon, t/\varepsilon^2\right)$$
$$+ \varepsilon \sum_{|\beta|=3} d_\beta \mathbf{h}_\beta(x/\varepsilon) \left[\partial_x^\beta \mathbf{m}_0 \times \mathbf{f}_v\right](x,t)\Psi\left(x/\varepsilon, t/\varepsilon^2\right) \Big).$$

Application of Lemma 4.7 thus yields that

$$(5.19) \qquad \varepsilon|III_3| = \left| \int_0^\eta \int_{\Omega_\mu} K_\eta^0 K_\mu \left(\mathcal{L}\mathbf{u}^\varepsilon \times \mathbf{v}^\varepsilon\right) dx dt \right|$$
$$\leq C \left(\|\mathbf{g}\|_{L^\infty} + \|\boldsymbol{\nabla}_y \mathbf{h}\|_{L^\infty} + \varepsilon\|\mathbf{h}\|_{L^\infty}\right)\varepsilon \left(\frac{\varepsilon^2}{\eta}\right)^{q_t+1}.$$

Finally, consider the last term in (5.15), $\mathbf{v}^\varepsilon \times \mathcal{L}\mathbf{v}^\varepsilon$, which can be bounded using Lemma 4.4 and the fact that $\mathbf{v}^\varepsilon$ decays exponentially in time. As we mostly consider spatial averaging only, we in the following suppress the time dependence of $\mathbf{v}$ for matters of brevity and write $\mathbf{v}(x,y)$ instead of $\mathbf{v}(x,y,t,\tau)$. Furthermore, we use the notation

$$(5.20) \qquad\qquad \mathcal{L}_{ab}\mathbf{v} := \boldsymbol{\nabla}_a \cdot (a(y)\boldsymbol{\nabla}_b \mathbf{v}).$$

Then we can write

$$\varepsilon^2 \mathbf{v}^\varepsilon \times \mathcal{L}\mathbf{v}^\varepsilon = \varepsilon^2 \mathbf{v}^\varepsilon \times \left[\mathcal{L}_{xx} + \varepsilon^{-1}(\mathcal{L}_{xy} + \mathcal{L}_{yx}) + \varepsilon^{-2}\mathcal{L}_{yy}\right]\mathbf{v}^\varepsilon$$
$$=: \mathbf{v}^\varepsilon \times \mathcal{L}_{yy}\mathbf{v}^\varepsilon + \varepsilon\hat{\mathbf{f}}_0^\varepsilon + \varepsilon^2 \hat{\mathbf{f}}_1^\varepsilon,$$

Consider first the average of $\mathbf{v} \times \mathcal{L}_{yy}\mathbf{v}$. Using integration by parts, we find that for any coordinate direction ℓ,

$$\int_0^1 \mathbf{v} \times \partial_{y_\ell}(a\partial_{y_\ell}\mathbf{v})dy_\ell = -\int_0^1 a\partial_{y_\ell}\mathbf{v} \times \partial_{y_\ell}\mathbf{v}dy_\ell = 0,$$

which implies that the average of $\mathbf{v} \times \mathcal{L}_{yy}\mathbf{v}$ with respect to y is zero. Moreover, we obtain using (1.3) and (3.11) that

$$\sup_{y\in Y} \|\mathbf{v}(\cdot,y) \times \mathcal{L}_{yy}\mathbf{v}(\cdot,y)\|_{W^{p_x+1,\infty}} \leq C\|\mathbf{v}\|_{H^{p_x+3,2}}\|\mathbf{v}\|_{H^{p_x+3,4}} \leq Ce^{-2\gamma t/\varepsilon^2}.$$

As $\mathbf{v}$ is periodic in y, spatial averaging according to Lemma 4.4 thus yields

$$\left| \int_{\Omega_\mu} K_\mu(\mathbf{v}^\varepsilon \times \mathcal{L}_{yy}\mathbf{v}^\varepsilon)dx \right| \leq Ce_{44}(\varepsilon,\mu)e^{-2\gamma t/\varepsilon^2}.$$

The averages of $\hat{\mathbf{f}}_0^\varepsilon$ and $\hat{\mathbf{f}}_1^\varepsilon$ can be bounded using (4.10) in Lemma 4.5. Consequently, it holds for $\varepsilon^2 III_4$ that

$$(5.21) \qquad \varepsilon^2 III_4 \leq \left| \int_0^\eta \int_{\Omega_\mu} K_\eta^0 K_\mu(\mathbf{v}^\varepsilon \times \mathcal{L}_{yy}\mathbf{v}^\varepsilon)dx dt \right| + \varepsilon\left|\bar{\mathcal{K}}_{\mu,\eta}\hat{\mathbf{f}}_0^\varepsilon\right| + \varepsilon^2\left|\bar{\mathcal{K}}_{\mu,\eta}\hat{\mathbf{f}}_1^\varepsilon\right|$$
$$\leq C\left(\varepsilon + \varepsilon^2 + e_{44}(\varepsilon,\mu)\int_0^\eta \left|K_\eta^0\right|e^{-2\gamma t/\varepsilon^2}dt\right) \leq C\left(\varepsilon + e_{44}(\varepsilon,\mu)\right).$$

Based on (5.15) together with (5.17), (5.18), (5.19) and (5.21), we thus obtain the estimate in Lemma 5.4. □

5.4. Remainder terms. In this section, we aim to bound the remainder errors r_{M1}, r_{M2} and r_{M3} that are introduced in the averaging processes of the models $(M1)-$ $(M3)$ when approximating $\mathbf{m}^\varepsilon$ by $\mathbf{m}^{\mathrm{app}}$ as shown in (5.1) - (5.3). To be able to formulate common results for these remainder terms, we introduce four different linear operators,

$$\mathcal{A}_1 \mathbf{u} := a^\varepsilon \boldsymbol{\nabla} \mathbf{u}, \qquad\qquad\qquad \mathcal{A}_2 \mathbf{u} := \mathcal{L} \mathbf{u},$$
$$\mathcal{A}_3 \mathbf{u} := \mathbf{m}^{\mathrm{app}} \times \mathcal{L} \mathbf{u}, \qquad\qquad \mathcal{A}_3 \mathbf{u} := \mathbf{u} \times \mathcal{L} \mathbf{m}^{\mathrm{app}},$$

where $\mathbf{u} \in H^2(\Omega)$ and $\mathbf{m}^{\mathrm{app}} := \mathbf{m}_0 + \varepsilon \mathbf{m}_1^\varepsilon$ as in the previous section. In terms of these operators, the remainder terms are

$$r_{M1} = \left| \bar{\mathcal{K}}_{\mu,\eta}(\mathcal{A}_1(\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}})) \right|,$$
$$r_{M2} = \left| \bar{\mathcal{K}}_{\mu,\eta}(\mathcal{A}_2(\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}})) \right|,$$
$$r_{M3} = \left| \bar{\mathcal{K}}_{\mu,\eta}(\mathcal{A}_3(\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}}) + \mathcal{A}_4(\mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}}) \right|.$$

To bound the remainder errors, we first consider the approximation $\tilde{\mathbf{m}}_J^\varepsilon$ as given by (3.3) and let $\mathbf{e}_J := \mathbf{m}^\varepsilon - \tilde{\mathbf{m}}_J^\varepsilon$. Note that by definition, $\mathbf{m}^{\mathrm{app}} = \tilde{\mathbf{m}}_1^\varepsilon$. It thus holds that

$$(5.22) \qquad\qquad \mathbf{m}^\varepsilon - \mathbf{m}^{\mathrm{app}} = \mathbf{e}_J + \sum_{j=2}^J \varepsilon^j \mathbf{m}_j^\varepsilon,$$

for some constant $J > 1$, where $\mathbf{m}_j^\varepsilon(x,t) := \mathbf{m}_j(x, x/\varepsilon, t, t/\varepsilon^2)$. As explained in Section 3, bounds for both H^q-norms of $\mathbf{e}_J$ as well as for the norms of $\mathbf{m}_j^\varepsilon$ have been proved in [17] and can be used to obtain estimates here. It is important to notice that the correctors $\mathbf{m}_j(x,y,t,\tau)$, $j > 0$, are periodic in y. We can therefore apply Lemma 4.4 to averages of these terms and their derivatives. The error $\mathbf{e}_J$, on the other hand, can in general not be assumed to be periodic and therefore has to be treated differently.

According to (4.9) in Lemma 4.5, it holds for the averages of the linear operators $\mathcal{A}_k$, $k = 1, ..., 4$ applied to $\mathbf{e}_J$ that

$$(5.23) \qquad \left| \bar{\mathcal{K}}_{\mu,\eta}(\mathcal{A}_k \mathbf{e}_J) \right| \leq \sup_{0 \leq t \leq \eta} \frac{1}{\mu^{d/2}} \|\mathcal{A}_k \mathbf{e}_J(\cdot, t)\|_{L^2} \leq \sup_{0 \leq t \leq \eta} \frac{1}{\varepsilon^{d/2}} \|\mathcal{A}_k \mathbf{e}_J(\cdot, t)\|_{L^2}.$$

Using the fact that by (3.8), for $0 \leq t \leq T^\varepsilon = \varepsilon^\sigma T$ with $1 < \sigma \leq 2$,

$$\|\mathbf{m}^{\mathrm{app}}\|_{W^{q,\infty}} = \|\tilde{\mathbf{m}}_1^\varepsilon\|_{W^{q,\infty}} \leq C\varepsilon, \qquad q \geq 1,$$

and the bound (3.9) for $\|\mathbf{e}_J\|_{H^q}$, the L^2-norms on the right-hand side in (5.23) can be bounded for $0 \leq t \leq T^\varepsilon$ as follows,

$$\|\mathcal{A}_1 \mathbf{e}_J(\cdot, t)\|_{L^2} \leq C \|\mathbf{e}_J(\cdot, t)\|_{H^1} \leq C\varepsilon^{\sigma_J + 1},$$
$$\|\mathcal{A}_2 \mathbf{e}_J(\cdot, t)\|_{L^2} \leq C \left(\varepsilon^{-1} \|\mathbf{e}_J(\cdot, t)\|_{H^1} + \|\mathbf{e}_J(\cdot, t)\|_{H^2} \right) \leq C\varepsilon^{\sigma_J},$$
$$\|\mathcal{A}_3 \mathbf{e}_J(\cdot, t)\|_{L^2} \leq C \|\mathbf{m}^{\mathrm{app}}\|_{W^{1,\infty}} \left(\varepsilon^{-1} \|\mathbf{e}_J(\cdot, t)\|_{H^1} + \|\mathbf{e}_J(\cdot, t)\|_{H^2} \right) \leq C\varepsilon^{\sigma_J},$$
$$\|\mathcal{A}_4 \mathbf{e}_J(\cdot, t)\|_{L^2} \leq C \|\mathbf{e}_J(\cdot, t)\|_{L^2} \left(\varepsilon^{-1} \|\mathbf{m}^{\mathrm{app}}\|_{W^{1,\infty}} + \|\mathbf{m}^{\mathrm{app}}\|_{W^{2,\infty}} \right) \leq C\varepsilon^{\sigma_J + 1},$$

where $\sigma_J := (\sigma - 1)(J - 1)$. We now choose $J \geq (1 + d/2)/(\sigma - 1) + 1$. Then $\sigma_J \geq 1 + d/2$ and since $\eta \leq T^\varepsilon$, it follows together with (5.23) that

$$(5.24) \qquad\qquad \left| \bar{\mathcal{K}}_{\mu,\eta}(\mathcal{A}_k \mathbf{e}_J) \right| \leq C\varepsilon, \qquad k = 1, ..., 4.$$

This provides a bound for the first part of the remainder errors. To also bound the second part, we use the L^∞-boundedness of $\bar{\mathcal{K}}_{\mu,\eta}$ as well as the periodicity of the correctors $\mathbf{m}_j$. To simplify these considerations, we split the operator $\mathcal{A}_3$,

$$\mathcal{A}_3\mathbf{u} = \mathbf{m}_0 \times \mathcal{L}\mathbf{u} + \varepsilon\mathbf{m}_1 \times \mathcal{L}\mathbf{u} =: \mathcal{A}_{31}\mathbf{u} + \mathcal{A}_{32}\mathbf{u}.$$

Note that by (3.7), we have for the considered T^ε and any $p, q > 0$

$$\tag{5.25} \sup_{0\leq t\leq T^\varepsilon} \|\mathbf{m}_j(\cdot,\cdot,t,t/\varepsilon^2)\|_{H^{p,q}} \leq C\varepsilon^{\min(0,2-j)}, \qquad j \geq 1.$$

Together with Lemma 5.1 in [17], this implies that

$$\sup_{0\leq t\leq T^\varepsilon} \|\mathbf{m}_j^\varepsilon(\cdot,t)\|_{W^{q,\infty}} \leq \sup_{0\leq t\leq T^\varepsilon} C\varepsilon^{-q}\|\mathbf{m}_j(\cdot,\cdot,t,t/\varepsilon^2)\|_{H^{q+2,q+2}} \leq C\varepsilon^{\min(0,2-j)-q}.$$

It therefore follows by (4.10) in Lemma 4.5 that for $j \geq 2$ and $0 \leq t \leq T^\varepsilon$,

$$\left|\bar{\mathcal{K}}_{\mu,\eta}\mathcal{A}_1\mathbf{m}_j^\varepsilon\right| \leq C\|\mathcal{A}_1\mathbf{m}_j^\varepsilon\|_{L^\infty} \leq C\|\mathbf{m}_j^\varepsilon\|_{W^{1,\infty}} \leq C\varepsilon^{1-j},$$
$$\left|\bar{\mathcal{K}}_{\mu,\eta}\mathcal{A}_{32}\mathbf{m}_j^\varepsilon\right| \leq C\|\mathcal{A}_{32}\mathbf{m}_j^\varepsilon\|_{L^\infty} \leq C\varepsilon\|\mathbf{m}_1^\varepsilon\|_{L^\infty}\left(\varepsilon^{-1}\|\mathbf{m}_j^\varepsilon\|_{W^{1,\infty}} + \|\mathbf{m}_j^\varepsilon\|_{W^{2,\infty}}\right) \leq C\varepsilon^{1-j},$$
$$\left|\bar{\mathcal{K}}_{\mu,\eta}\mathcal{A}_4\mathbf{m}_j^\varepsilon\right| \leq C\|\mathcal{A}_4\mathbf{m}_j^\varepsilon\|_{L^\infty}$$
$$\leq C\|\mathbf{m}_j^\varepsilon\|_{L^\infty}\left(\varepsilon^{-1}\|\mathbf{m}^{\mathrm{app}}\|_{W^{1,\infty}} + \|\mathbf{m}^{\mathrm{app}}\|_{W^{2,\infty}}\right) \leq C\varepsilon^{1-j},$$

To bound the remaining terms, $|\bar{\mathcal{K}}_{\mu,\eta}\mathcal{A}_2\mathbf{m}_j|$ and $|\bar{\mathcal{K}}_{\mu,\eta}\mathcal{A}_{31}\mathbf{m}_j|$, we furthermore have to exploit the periodicity of the correctors. We proceed in a similar way as for III_4 in the previous section, applying Lemma 4.4 and again using the notation given in (5.20). Suppressing time dependence, let

$$\mathbf{f}_j(x,y) := \left(\mathcal{L}_{xx} + \varepsilon^{-1}(L_{xy} + L_{yx}) + \varepsilon^{-2}\mathcal{L}_{yy}\right)\mathbf{m}_j(x,y)$$

As the Y-average of $\mathcal{L}_{yy}\mathbf{m}_j$ is zero, we then find using (5.25) that for $j \geq 2$,

$$\left|\fint_Y \mathbf{f}_j(0,y)dy\right| \leq C\left(\|\mathbf{m}_j\|_{H^{4,0}} + \varepsilon^{-1}\|\mathbf{m}_j\|_{H^{3,1}}\right) \leq C\varepsilon^{1-j}$$

and

$$\sup_{y\in Y}\|\mathbf{f}_j(\cdot,y)\|_{W^{p_x+1,\infty}}$$
$$\leq C\left(\|\mathbf{m}_j\|_{H^{p_x+5,2}} + \varepsilon^{-1}\|\mathbf{m}_j\|_{H^{p_x+4,3}} + \varepsilon^{-2}\|\mathbf{m}_j\|_{H^{p_x+3,4}}\right) \leq C\varepsilon^{-j}.$$

Hence, the spatial average of $\mathcal{A}_2\mathbf{m}_j$ is according to Lemma 4.4 bounded as follows,

$$\int_{\Omega_\mu} K_\mu(x)\mathcal{A}_2\mathbf{m}_j(x,x/\varepsilon)dx \leq \left|\fint_Y \mathbf{f}_j(0,y)dy\right| + C\sup_{y\in Y}\|\mathbf{f}_j\|_{W^{p_x+1,\infty}}e_{44}(\varepsilon,\mu)$$
$$\leq C\varepsilon^{-j}\left(\varepsilon + e_{44}(\varepsilon,\mu)\right),$$

and similarly,

$$\int_{\Omega_\mu} K_\mu\mathcal{A}_{31}\mathbf{m}_j dx \leq \left|\mathbf{m}_0(0) \times \fint_Y \mathbf{f}_j(0,y)dy\right| + C\sup_{y\in Y}\|\mathbf{m}_0 \times \mathbf{f}_j\|_{W^{p_x+1,\infty}}e_{44}(\varepsilon,\mu)$$
$$\leq C\varepsilon^{-j}\left(\varepsilon + e_{44}(\varepsilon,\mu)\right).$$

We can therefore conclude that for $\ell \in \{2, 31\}$,

$$\left| \bar{\mathcal{K}}_{\mu,\eta} \mathcal{A}_\ell \mathbf{m}_j \right| \leq \int_0^\eta |K_\eta^0| \left| \int_{\Omega_\mu} K_\mu \mathcal{A}_\ell \mathbf{m}_j dx \right| dt \leq C\varepsilon^{-j} \left(\varepsilon + e_{44}(\varepsilon, \mu) \right).$$

Overall, we finally obtain that when choosing $J \geq (1 + d/2)/(\sigma - 1) + 1$,

$$r_{M_1} \leq \left| \bar{\mathcal{K}}_{\mu,\eta} \mathcal{A}_1 \mathbf{e}_J \right| + \sum_{j=2}^J \varepsilon^j \left| \bar{\mathcal{K}}_{\mu,\eta} \mathcal{A}_1 \mathbf{m}_j \right| \leq C\varepsilon,$$

$$r_{M_2} \leq \left| \bar{\mathcal{K}}_{\mu,\eta} \mathcal{A}_2 \mathbf{e}_J \right| + \sum_{j=2}^J \varepsilon^j \left| \bar{\mathcal{K}}_{\mu,\eta} \mathcal{A}_2 \mathbf{m}_j \right| \leq C\varepsilon + e_{44}(\varepsilon, \mu),$$

$$r_{M_3} \leq \left| \bar{\mathcal{K}}_{\mu,\eta} (\mathcal{A}_3 + \mathcal{A}_4) \mathbf{e}_J \right| + \sum_{j=2}^J \varepsilon^j \left| \bar{\mathcal{K}}_{\mu,\eta} (\mathcal{A}_{31} + \mathcal{A}_{32} + \mathcal{A}_4) \mathbf{m}_j \right| \leq C\varepsilon + e_{44}(\varepsilon, \mu).$$

This completes the proof of Theorem 5.1.

6. Numerical results. In this section, we present numerical examples showing the convergence of the averaged flux, field and torque as specified in the models (M1) - (M3) to the corresponding homogenized quantities in both one and two space dimensions. We provide evidence for the estimates given in Theorem 5.1 and show which of the terms appearing there seem to be dominating in practice.

6.1. One-dimensional examples. In one space dimension, we consider the periodic material coefficient

$$(6.1) \qquad a^\varepsilon(x) = a(x/\varepsilon), \qquad \text{where} \qquad a(y) = 1 + 0.5\sin(2\pi y) + 0.5\sin(4\pi y).$$

The corresponding homogenized coefficient, given by

$$a^H = \left(\int_0^1 (a(y))^{-1} dy \right)^{-1},$$

is computed numerically. To create a better understanding of the problem, Figure 1 shows the x-component of the solution $\mathbf{m}^\varepsilon$ to an example problem in time and space as well as the difference between $\mathbf{m}^\varepsilon$ and the solution to the corresponding homogenized equation, $\mathbf{m}_0$. One can observe that $\mathbf{m}^\varepsilon$ oscillates in both time and space initially, but as t increases the temporal oscillations are damped away and only the spatial ones remain. The oscillations have significantly smaller magnitude than the solution and appear to have zero average.

We then consider the approximation errors for the three models (M1) - (M3) and compare the observed behavior with the theoretical bound according to Theorem 5.1,

$$(6.2) \quad E_{Mi} \leq C \left(\varepsilon + \left(\frac{\varepsilon}{\mu} \right)^{q_x+2} + \mu^{p_x+1} + \eta^{p_t+1} + \frac{1}{\mu^{\delta_i}} \left(\frac{\varepsilon^2}{\eta} \right)^{q_t+1} \right), \quad i = 1, 2, 3,$$

where $\delta_i = 0, i = 1$ and $\delta_i = 1, i = 2, 3$. In Figure 2, the approximation errors for varying ε is shown. In Figure 2a, the damping constant is set to $\alpha = 0.01$ and in Figure 2b we have $\alpha = 0.1$. For all three models, the errors initially decrease rapidly with ε. In this regime, the error appears to be dominated by the ε^2/η term in (6.2).
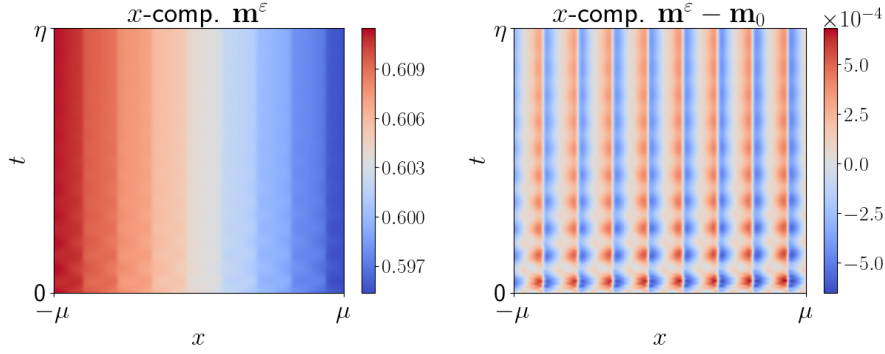
Fig. 1: Left: Development of x-component of solution $\mathbf{m}^\varepsilon$ to (2.1) with a^ε given by (6.1) in space and time when $\varepsilon = 1/140$ and $\alpha = 0.05$. Spatial domain $[-\mu, \mu]$ where $\mu = 0.03$, time interval $t \in [0, \eta]$ with $\eta = 1.5 \cdot 10^{-4}$. Right: difference between that solution $\mathbf{m}^\varepsilon$ and corresponding homogenized solution $\mathbf{m}_0$.
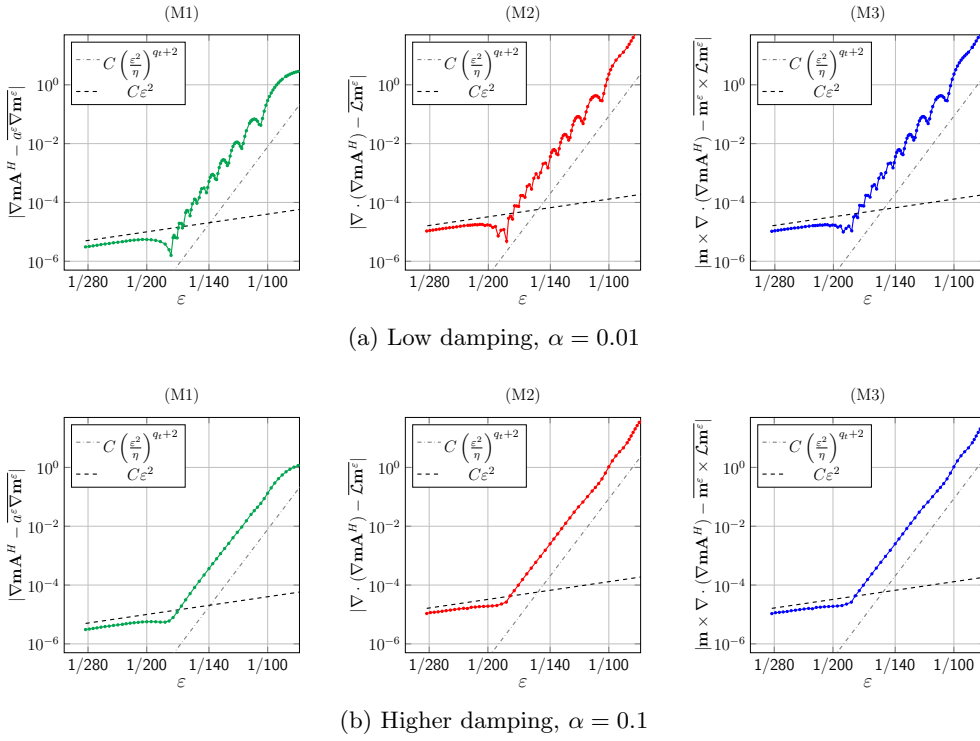

(a) Low damping, $\alpha = 0.01$

(b) Higher damping, $\alpha = 0.1$

Fig. 2: Approximation errors in (M1)-(M3) for varying ε with kernel parameters $p_x = 5, q_x = 7, p_t = 5, q_t = 7$ and averaging domain sizes $\eta = 1.5 \cdot 10^{-4}$ and $\mu = 0.03$.

For smaller ε, the errors are proportional to ε^2. Note that this is smaller than the convergence rate of ε suggested by (6.2). In (M1) the error is somewhat lower than

for (M2) and (M3). The latter two models result in very similar error behavior.

When comparing Figure 2a and Figure 2b, one can moreover observe that choosing a lower damping parameter α results in more oscillatory errors. However, the overall error behavior is very similar for both α-values. This is a property that holds for all the examples considered here. We therefore choose to show plots for higher values of α in several of the subsequent examples to reduce oscillations and make it easier to distinguish the different curves.

Next, we consider the influence of the kernel parameters q_x and q_t on the error decay. As shown in Figure 3, the choice of q_x does not influence the error behavior, while different values of q_t result in different slopes of the initial error decay. This again shows that the error from time averaging initially dominates in Figure 2 and Figure 3. In particular, it is proportional to $(\varepsilon^2/\eta)^{q_t+2}$ until it reaches the ε^2 threshold. This is slightly better than $(\varepsilon^2/\eta)^{q_t+1}$ as given in Theorem 5.1.
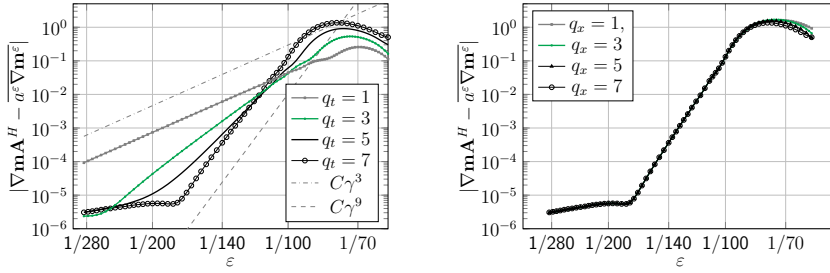
Fig. 3: Approximation error for (M1) for varying ε with $\alpha = 0.1$ and kernel parameters $p_x = 5, q_x = 7, p_t = 5$ and $q_t = 7$, if not explicitly stated otherwise in the plot. Averaging domain sizes $\eta = 1.5 \cdot 10^{-4}$ and $\mu = 0.03$. For reference lines, $\gamma := \varepsilon^2/\eta$.

We continue by examining the influence of the parameters p_x and p_t and the contribution of the terms $C\mu^{p_x+1}$ and $C\eta^{p_t+1}$ to the error. As shown in Figure 4, we find that low choices of both p_x and p_t result in a constant error for low ε, corresponding to $C\mu^{p_x+1}$ or $C\eta^{p_t+1}$, respectively. For larger p_x and p_t, these terms are presumably smaller than ε^2, in the range considered. One can furthermore observe that the choice of p_x does not seem to influence the error otherwise, while the initial convergence happens at different ε-values when varying p_t.
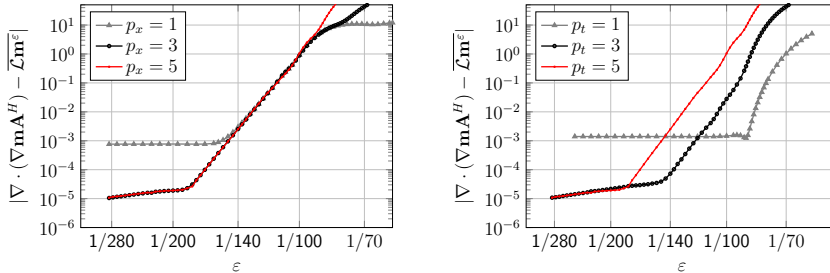

Fig. 4: Approximation error for the model (M2) for varying ε with kernel parameters $p_x = 5, q_x = 7, p_t = 5$ and $q_t = 7$, if not explicitly stated otherwise in the plot. Averaging domain sizes $\eta = 1.5 \cdot 10^{-4}$ and $\mu = 0.03$ and damping $\alpha = 0.1$.

Finally, we investigate the influence of the choice of box sizes μ and η on the error given a fixed value $\varepsilon = 1/140$. In Figure 5, it is shown that when increasing μ from a small value, there is some initial decrease in the error due to the reduction in $(\varepsilon/\mu)^{q_x+2}$ before it takes a constant value, due to the fact that the other terms in (6.2) dominate. At some point, depending on p_x, the error starts increasing again since the term μ^{p_x+1} starts dominating the error. When varying η, we have a similar behavior. However, increasing η results in a much larger initial error reduction, given that p_t is chosen large enough. The slopes of this decrease depend on q_t. Once η is larger than a certain threshold, the error takes a constant value. When $p_t = 1$, the error only decreases initially and then starts increasing again due to the term η^{p_t+1}.
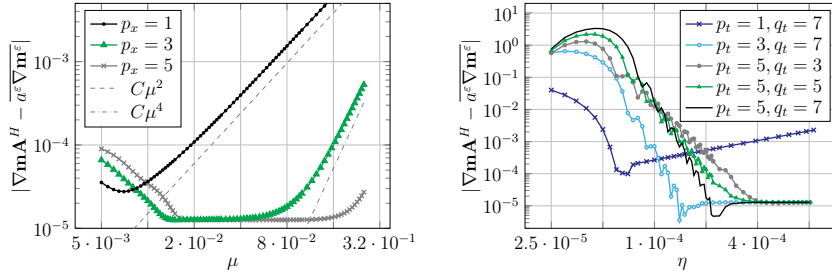
Fig. 5: Influence of the box sizes η and μ on the error in (M1) for fixed $\varepsilon = 1/140$. Kernel parameters $p_x = 5, q_x = 7, p_t = 5, q_t = 7$ if not explicitly stated otherwise in the plot. Damping $\alpha = 0.01$. When varying η, $\mu = 0.03$ and when varying μ, $\eta = 1.5 \cdot 10^{-4}$.

Overall, we can conclude that for the 1D example problem, the error is considerably more affected by temporal than spatial averaging. The estimate in Theorem 5.1 matches conceptually well with the observed behavior but is slightly too pessimistic.

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

$$(6.3) \qquad a^\varepsilon(x) = e^{\cos(2\pi x/\varepsilon)\cos(2\pi x)}.$$

The corresponding homogenized coefficient at $x = 0$ is computed by fixing the slow variable and proceeding as in the periodic case. The resulting approximation errors for the models (M1)-(M3) are depicted in Figure 6. Even though locally periodic coefficients are not covered by Theorem 5.1, nor by the general homogenization theory in [17], the observed errors behave similar to the periodic case, with a fast initial decrease of the error as ε is decreased and then convergence proportional to ε^2. The errors for field and torque model are again very similar. The flux error is somewhat lower.

6.2. Two-dimensional examples. We here consider a periodic problem where the material coefficient is chosen to be

$$a(x_1, x_2) = \frac{1}{2} + \left(\frac{1}{2} + \frac{1}{4}\sin(2\pi x_1)\right)\left(\frac{1}{2} + \frac{1}{4}\sin(2\pi x_2)\right)$$
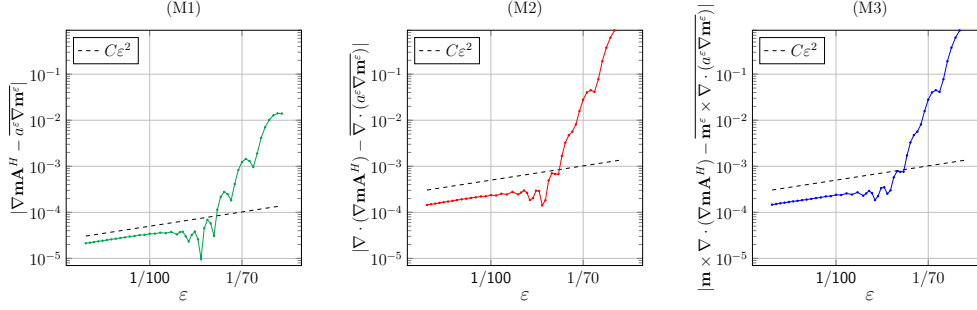$$+ \frac{1}{4}\cos(2\pi(x_1 - x_2)) + \frac{1}{2}\sin(2\pi x_1).$$

Fig. 6: Approximation errors in (M1)-(M3) with locally periodic material coefficient $a^\varepsilon(x)$ as in (6.3) for varying ε. Here $\alpha = 0.01$, the kernel parameters are $p_x = 3, q_x = 7, p_t = 3, q_t = 7$ and averaging domain sizes $\eta = 2.1 \cdot 10^{-4}$ and $\mu = 0.03$.

Solving the cell problem (1.7) numerically, the corresponding homogenized coefficient is computed to be

$$\mathbf{A}^H = \begin{bmatrix} 0.61720765 & 0.02618130 \\ 0.02618130 & 0.71523722 \end{bmatrix},$$

a full matrix with two different diagonal elements. The upscaling errors when varying ε in (M1), (M2) and (M3) for an example problem with this material coefficient is shown in Figure 7.

One can observe a similar behavior as for the 1D problem. However, note that the error in (M1) is considerably lower than for (M2) and (M3) in this example. In (M1) and (M2), we again observe convergence proportional to ε^2 for low values of ε instead of ε as suggested by Theorem 5.1. However, the error in (M3) with low damping, $\alpha = 0.01$, decays only proportionally to ε rather than ε^2. We suspect that this is related to the term εIII_4 in the analysis of the error in (M3), (5.21), the term taking the interaction of fast oscillations in time with each other into account. With higher α, the temporal oscillations get damped away faster and we do not observe that behavior.
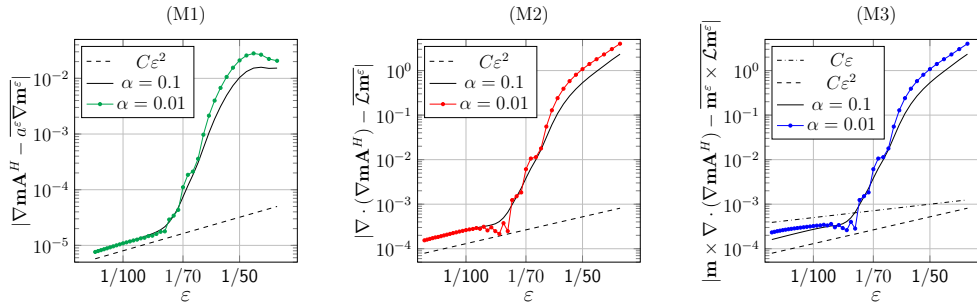

Fig. 7: Approximation error in (M1)-(M3) two dimensions for varying ε, where $\mu = 0.06, \eta = 0.0003$ and the kernel parameters are $p_x = 5, q_x = 7, p_t = 3$ and $q_t = 7$.

Apart from this observation for low ε, the errors in (M2) and (M3) behave very similar when varying the parameters in the model. We therefore focus on comparing

(M1) and (M2) in the following. The influence of the kernel parameters is similar to the 1D problem. As shown in Figure 8, choosing low p_x or p_t results in constant error when decreasing ε, corresponding to $C\mu^{p_x+1}$ or $C\eta^{p_t+1}$, respectively. The parameter q_t determines the speed of the initial decay. However, in contrast to the 1D case it is harder to specifically determine the slopes in this example.
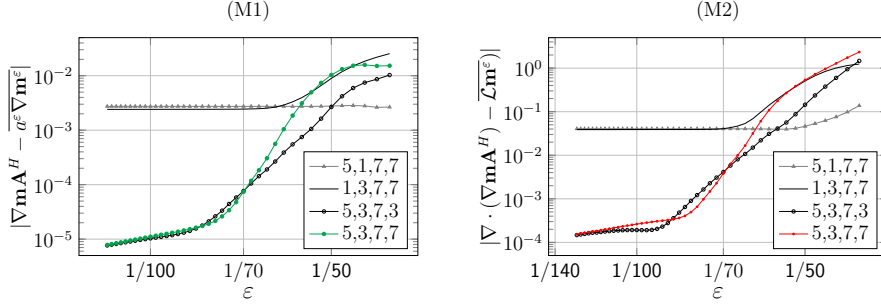

Fig. 8: Approximation error in (M1) and (M2) when varying ε. Damping $\alpha = 0.1$, spatial averaging parameter $\mu = 0.06$ and averaging time $\eta = 0.0003$. Kernel parameters according to legend, in order p_x, p_t, q_x, q_t.

In Figure 9, the error in (M1) and (M2) when varying η is shown for two different values of ε, similar to Figure 5, right, in the 1D case. When choosing low η, the errors are high but decrease rapidly as η increases. From $\eta \approx 2\varepsilon^2$ the error stays at a constant level.
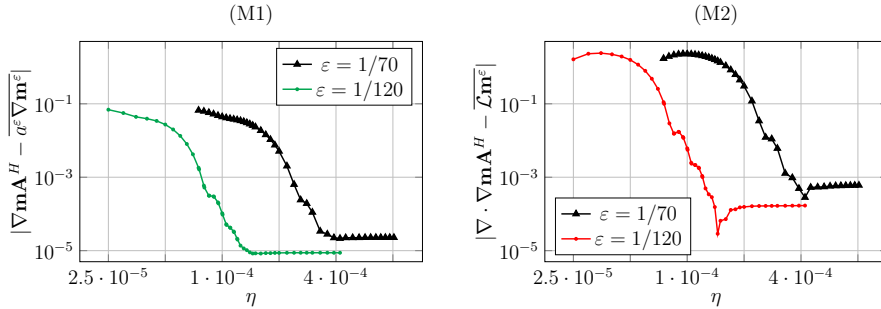

Fig. 9: Approximation error in (M1) and (M2) when increasing η. Damping $\alpha = 0.01$, kernel parameters $p_x = 5, q_x = 7, p_t = 3, q_t = 7$ and $\mu = 0.06$.

Finally, we investigate the influence of μ as shown in Figure 10. We can observe rapidly decreasing errors until $\mu \approx 3\varepsilon$, then the errors are almost constant. In contrast to the 1D case, shown in Figure 5, left, the choice of μ has a significant impact on the error in this example. In particular, in (M1) the magnitude of the error is determined by η and μ equally. In case of (M2), η still has a somewhat larger impact than μ.

Finally, we consider a two-dimensional locally periodic example. In this case, the material coefficient is

$$(6.4) \qquad a^\varepsilon(x_1, x_2) = 1.1 + 0.5\sin(2\pi x_1/\varepsilon) + 0.5\sin(2\pi x_2)\cos(2\pi x_1/\varepsilon)$$
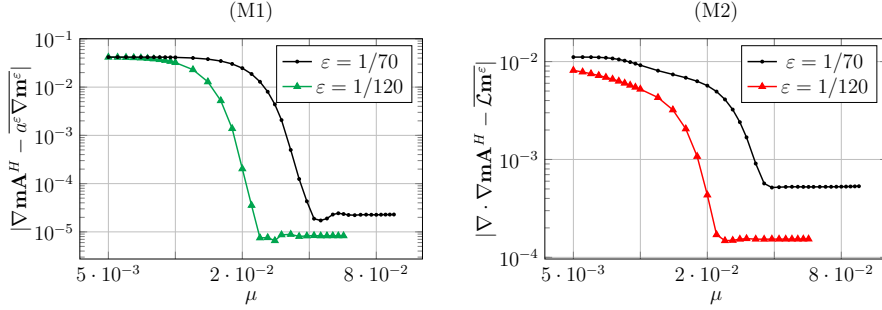
Fig. 10: Approximation error in (M1) and (M2) when increasing the spatial averaging size μ. Damping $\alpha = 0.01$ and kernel parameters $p_x = 5, q_x = 7, p_t = 3, q_t = 7$. For $\varepsilon = 1/70$, final time $\eta = 4.5 \cdot 10^{-4}$ and for $\varepsilon = 1/120$, $\eta = 2 \cdot 10^{-4}$.

As in the one-dimensional case, the reference solution is computed for a fixed slow variable value, $x_2 = 0$. The resulting approximation errors are shown in Figure 11. Also in this case, the errors initially converge rapidly as ε decreases. However, in contrast to the periodic case, for smaller ε-values, we here observe convergence proportional to ε for all three models. This also is a difference compared to the one-dimensional locally-periodic example as shown in Figure 6, where the errors in that regime where proportional to ε^2 as in the periodic case.

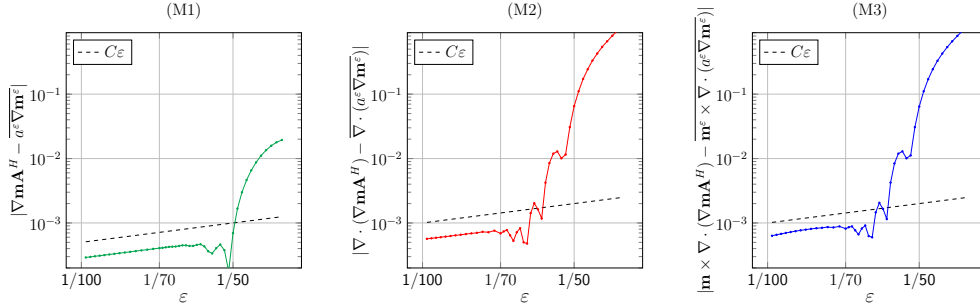Overall, the considered models still seem to be valid in the locally periodic case.

Fig. 11: Approximation errors in (M1)-(M3) with locally periodic material coefficient $a^\varepsilon(x)$ as in (6.4) for varying ε and $\alpha = 0.01$. Kernel parameters $p_x = 3, q_x = 7, p_t = 3, q_t = 7$ and averaging domain sizes $\eta = 4.1 \cdot 10^{-4}$ and $\mu = 0.06$.

REFERENCES

[1] A. ABDULLE, E. WEINAN, B. ENGQUIST, AND E. VANDEN-EIJNDEN, *The Heterogeneous Multiscale Method*, Acta Numerica, 21 (2012), pp. 1–87.
[2] F. ALOUGES, A. DE BOUARD, B. MERLET, AND L. NICOLAS, *Stochastic homogenization of the Landau-Lifshitz-Gilbert equation*, Stochastics and Partial Differential Equations: Analysis and Computations, (2021), pp. 1–30.

[3] D. ARJMAND, S. ENGBLOM, AND G. KREISS, *Temporal Upscaling in Micromagnetism via Heterogeneous Multiscale Methods*, J. Comput. Appl. Math., 345 (2019), pp. 99–113.

[4] D. ARJMAND, G. KREISS, AND M. POLUEKTOV, *Atomistic-continuum multiscale modeling of magnetization dynamics at non-zero temperature*, Adv. Comput. Math, 44 (2018), pp. 1119–1151.

[5] D. ARJMAND AND O. RUNBORG, *Analysis of heterogeneous multiscale methods for long time wave propagation problems*, Multiscale Modeling & Simulation, 12 (2014), pp. 1135–1166.

[6] D. ARJMAND AND O. RUNBORG, *A time dependent approach for removing the cell boundary error in elliptic homogenization problems*, Journal of Computational Physics, 314 (2016), pp. 206–227.

[7] A. BENSOUSSAN, J. LIONS, AND G. PAPANICOLAOU, *Asymptotic analysis for periodic structures*, North Holland, 1978.

[8] C. CHOQUET, M. MOUMNI, AND M. TILIOUA, *Homogenization of the Landau-Lifshitz-Gilbert equation in a contrasted composite medium*, Discrete & Continuous Dynamical Systems-S, 11 (2018), p. 35.

[9] D. CIORANESCU AND P. DONATO, *An introduction to homogenization*, Oxford University Press, 1999.

[10] B. ENGQUIST, H. HOLST, AND O. RUNBORG, *Multi-scale methods for wave propagation in heterogeneous media*, arXiv preprint arXiv:0911.2638, (2009).

[11] B. ENGQUIST, H. HOLST, AND O. RUNBORG, *Analysis of HMM for one dimensional wave propagation problems over long time*, arXiv preprint arXiv:1111.2541, (2011).

[12] B. ENGQUIST AND Y.-H. TSAI, *Heterogeneous multiscale methods for stiff ordinary differential equations*, Mathematics of computation, 74 (2005), pp. 1707–1742.

[13] K. HAMDACHE, *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.

[14] I. G. KEVREKIDIS, C. W. GEAR, J. HYMAN, P. G. KEVEKIDIS, AND O. RUNBORG, *Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level tasks.*, Comm. Math. Sci., (2003), pp. 715–762.

[15] M. G. KREIN AND M. A. RUTHMAN, *Linear operators that leave invariant a cone in a Banach space*, Usp. Mat. Nauk., (1948).

[16] L. LANDAU AND E. LIFSHITZ, *On the theory of dispersion of magnetic permeability in ferromagnetic bodies*, Phys. Z. Sowjet., 8 (1935), pp. 153–168.

[17] L. LEITENMAIER AND O. RUNBORG, *Homogenization of the Landau-Lifshitz equation*, arXiv preprint arXiv:2012.12567, (2020).

[18] L. LEITENMAIER AND O. RUNBORG, *Heterogeneous Multiscale Methods for the Landau-Lifshitz equation*, arXiv preprint arXiv:2108.09463, (2021).

[19] M. J. D. POWELL ET AL., *Approximation theory and methods*, Cambridge University Press, 1981.

[20] K. SANTUGINI-REPIQUET, *Homogenization of ferromagnetic multilayers in the presence of surface energies*, ESAIM: Control, Optimisation and Calculus of Variations, 13 (2007), pp. 305–330.

[21] E. WEINAN, B. ENGQUIST, ET AL., *The Heterognous Multiscale Methods*, Communications in Mathematical Sciences, 1 (2003), pp. 87–132.

[22] E. WEINAN, B. ENGQUIST, X. LI, W. REN, AND E. VANDEN-EIJNDEN, *Heterogeneous Multiscale Methods: a review*, Communications in computational physics, 2 (2007), pp. 367–450.