NETWORKS AND HETEROGENEOUS MEDIA
©American Institute of Mathematical Sciences
Volume **5**, Number **4**, December **2010**

doi:10.3934/nhm.2010.5.711

pp. **711–744**

THE HETEROGENEOUS MULTISCALE FINITE ELEMENT METHOD FOR ADVECTION-DIFFUSION PROBLEMS WITH RAPIDLY OSCILLATING COEFFICIENTS AND LARGE EXPECTED DRIFT

Patrick Henning and Mario Ohlberger

Institut für Numerische und Angewandte Mathematik
Fachbereich Mathematik und Informatik der Universität Münster
Einsteinstrasse 62, 48149 Münster, Germany

(Communicated by Alfio Quarteroni)

Abstract. This contribution is concerned with the formulation of a heterogeneous multiscale finite elements method (HMM) for solving linear advection-diffusion problems with rapidly oscillating coefficient functions and a large expected drift. We show that, in the case of periodic coefficient functions, this approach is equivalent to a discretization of the two-scale homogenized equation by means of a Discontinuous Galerkin Time Stepping Method with quadrature. We then derive an optimal order a-priori error estimate for this version of the HMM and finally provide numerical experiments to validate the method.

1. Introduction.

In this contribution we are concerned with the numerical analysis of a numerical multiscale finite element method for advection-diffusion problems where the coefficient functions have rapid oscillations within the space variable and where a large macroscopic drift may be expected. This means that we treat equations of the type

$$k^\epsilon \partial_t u^\epsilon - \nabla \cdot (A^\epsilon \nabla u^\epsilon) + \epsilon^{-1} b^\epsilon \cdot \nabla u^\epsilon = 0 \text{ in } \mathbb{R}^d \times (0, \bar{T}), \tag{1}$$
$$u^\epsilon(0, \cdot) = v_0 \text{ in } \mathbb{R}^d,$$

with a very small parameter ϵ that should be regarded as a measure for the degree of fineness of the problem. For the time-dependent coefficients we assume that they primarily contain microscopic oscillations, whereas the behaviour on the macro-scale is constant. Moreover, the velocity field b^ϵ is assumed to be divergence-free. Even though, the case where the coefficients are also allowed to vary on the macro-scale is not part of this work, a generalization to this situation is possible, but yields several difficulties. These difficulties already arise in the homogenization theory for such problems, where a so-called exponential spectral problem is required to formulate the homogenized equation (see Allaire and Orive [5]). We therefore postpone the analysis of the general case to future work. We note that problem (1) also covers the

2000 *Mathematics Subject Classification.* Primary: 35K15, 35B45, 65N30.

Key words and phrases. Advection-diffusion equation, HMM, multiscale methods, Finite Element scheme, error estimate.

This work was supported by the Bundesministerium für Bildung und Forschung under the contract number 03OMPAF1.

treatment of advection-diffusion-reaction problems, if the corresponding coefficients are periodic and independent of t, i.e. if the problem with additional reaction term is of the following kind:

$$\partial_t \tilde{u^\epsilon} - \nabla \cdot \left(\tilde{A}\left(\frac{x}{\epsilon}\right) \nabla \tilde{u^\epsilon} \right) + \epsilon^{-1} \tilde{b}\left(\frac{x}{\epsilon}\right) \cdot \nabla \tilde{u^\epsilon} + \epsilon^{-2} \tilde{c}\left(\frac{x}{\epsilon}\right) \tilde{u^\epsilon} = 0 \text{ in } \mathbb{R}^d \times (0, \bar{T}), \quad (2)$$
$$\tilde{u^\epsilon}(0, \cdot) = \tilde{v}_0 \text{ in } \mathbb{R}^d.$$

Here, the assumption that $\nabla \cdot \tilde{b} = 0$ is not needed. The matching was treated by Allaire and Raphael [6, 7] in 2007, who show that, by means of certain spectral cell problems, equation (2) can be transformed to a simple advection-diffusion problem with a divergence-free velocity field b. The transformation itself can be easily calculated and results among others in the additional coefficient function k.

The scaling of the convective term with $\frac{1}{\epsilon}$ in equation (1) refers to a large Péclet number (see for instance [9]). Hence, in this contribution we are interested in the advection dominated case.

Equations of type (1) have a variety of applications such as reservoir displacement problems, the modeling of semi-conductor devices, polymer chemistry and especially models for transport of solutes in groundwater and surface water. Here, we may for instance look at the modeling of groundwater pollution, where the interest is to determine the concentration of the contaminant in the water. In particular when the water content takes values that are close to saturation, large Péclet numbers will occur. This is due to the fact that saturation implies that the diffusion process has only a minor influence and the flow is primarily caused by gravity.

The problem with the numerical treatment of this type of equations, is the incredibly fine micro-structure, which also may have an important influence on the coarse scale properties. To compute reliable numerical approximations with a standard method for parabolic equations, this structure must be well resolved. Since this results in an intractable computational demand, we are interested in reducing the high complexity of such problems by formulating suitable methods that get by without global fine-scale calculations. In particular we are concerned with approaches that allow us to remain independent of the fineness parameter ϵ. Even if it is getting smaller, the computational demand remains constant.

There are several approaches dealing with that problem, which is why we give a small survey on the different techniques. Since some of the methods, treating elliptic homogenization problems, have not yet been adapted to parabolic equations, we also include the stationary case in our overview. A first example for a method to solve fine-scale problems, is the so called multiscale finite element method developed by Hou et al. This method is applicable to heterogeneous composite materials as well as to porous media. The elliptic case was treated in [20, 21], whereas the applications to two phase flow in porous media were observed in [14]. Multiscale methods for solving parabolic equations with continuum spatial scales and heterogeneous coefficients were discussed by Jiang, Efendiev and Ginting [22]. A projection framework for multiscale methods for the elliptic case is presented by Nolen, Papanicolaou and Pironneau [27]. Another possibility of finding a proper approximation to the solution u^ϵ of an elliptic homogenization problem, could be realized by means of a two-scale finite element method, such as proposed by Schwab and Matache [24, 25, 30]. Here it is assumed that the coefficients are periodically oscillating. Therefore, the method makes explicit use of the so-called two-scale homogenized equation that is equivalent to the standard homogenized problem. In [19], Hoang and Schwab introduce a two-scale FEM which is realized by a discretization of this formulation on sparse grids. In a contribution of Arbogast et al [8], a multiscale mortar mixed finite element discretizations for second order elliptic equations is treated. Here an overall domain Ω is subdivided into coarse elements, the subdomains, on which the original problem is posed. These subdomains are discretized on a very fine grid scale and are stringed together by a low degree-of-freedom mortar space. For a 'divide-and-conquer' spatial and temporal multiscale method for transient advection-diffusion-reaction equations, see the work of Gravemeier and Wall [15]. Adaptive algorithms for stationary fine-scale problems were developed by Oden and Vemaganti [28, 32]. These algorithms determine a number of cells in which the error between the homogenized solution and the exact solution is still too large. Locally on these cells, a fine-scale problem is solved whose solution is added as a perturbation to the homogenized solution. The local error indicators are measured in a quantity of interest, which could be a norm concerning the physical background. One further method to treat fine-scale problems is the heterogeneous multiscale finite element method (HMM), as shall be discussed in this paper. Initially introduced in 2003 by E and Engquist [10, 11, 12], the HMM is based on a standard finite element approach, whereas the evaluation of the corresponding discrete bilinear form is achieved by means of solving local cell problems in quadrature points. This method is not restricted to the case of periodicity. The HMM for elliptic problems on non-perforated domains was treated in contributions of E, Ming and Zhang [13], Abdulle and Schwab [4] and Ohlberger [29] and the perforated case by Henning and Ohlberger [18]. The parabolic case (HMM) was observed by Abdulle and E [3] and Ming and Zhang [26]. In another work of Abdulle [1], an algorithm for solving advection-diffusion problems is presented, where the HMM is combined with an Orthogonal Runge-Kutta Chebyshev (ROCK) method, in order to get an efficient resolution of the micro-structure.

Among others, a-priori results concerning HMM were achieved in [1, 2, 10] and [13]. The associated proofs, however, made direct use of the local problems belonging to the method, so that these approaches are not applicable to a further a-posteriori theory. To avoid this problem Ohlberger [29] reformulates the HMM into a discrete two-scale equation in order to compare this reformulation with the corresponding two-scale homogenized problem. On this basis a-posteriori estimates for the elliptic problem could be shown by Ohlberger for the case of domains without inclusions [29] and by Henning and Ohlberger for the case of a perforated domain [18].

The goal of this paper is an original formulation of the HMM for advection-diffusion problems with rapidly oscillating coefficient functions and a large expected drift. Since the large drift is a result of the microscopic behaviour, we integrate this heuristic ansatz into our approach. A detailed motivation behind the method will be given. For the case of periodic coefficient functions, we will show that our method is equivalent to a discretization of the two-scale homogenized equation by means of a Discontinuous Galerkin Time Stepping Method. Using this technique of a reformulation (see also [29] and [18]), the heterogeneous multiscale method is put into a variational framework, which simplifies the analysis. In this paper we focus on a-priori error estimates, whereas a-posteriori error estimates will be considered in a forthcoming work [16].

The article is structured into four main parts. Section 2 introduces some general assumptions and recalls several important analytic results of the homogenization theory. The next part is concerned with the derivation of the HMM for advection-diffusion problems and its reformulation under certain circumstances. In the following section the a-priori error estimate is derived, using this reformulated version. In the last part, we state two numerical model problems to show the applicability and efficiency of our method.

2. **General assumptions and analytic results.** In this section we are dealing with the periodic setting and the homogenization of equation (1). Moreover, we introduce all the definitions and notations that are required to formulate the heterogeneous multiscale finite element method.

2.1. **The continuous setting and a homogenization result.** This subsection is covering the treatment of the following linear advection-diffusion problem with rapidly oscillating coefficient functions and a large expected drift: find $u^\epsilon \in H^1(0,\bar{T};H^1(\mathbb{R}^d))$ with

$$\begin{aligned} &\int_0^{\bar{T}}\int_{\mathbb{R}^d} k\left(t,\frac{x}{\epsilon}\right)\partial_t u^\epsilon(t,x)\Phi(t,x) + A\left(t,\frac{x}{\epsilon}\right)\nabla u^\epsilon(t,x)\cdot\nabla\Phi(t,x)\,dx\,dt \qquad (3)\\ &\quad + \int_0^{\bar{T}}\int_{\mathbb{R}^d}\epsilon^{-1} b\Big(t,\frac{x}{\epsilon}\Big)\cdot\nabla u^\epsilon(t,x)\Phi(t,x)\,dx\,dt = 0 \quad \forall\Phi\in H^1(0,\bar{T},H^1(\mathbb{R}^d))\end{aligned}$$

and $u^\epsilon(0,\cdot) = v_0$. For our purposes, it is sufficient to assume that the coefficient functions fulfill some regularity and ellipticity properties, such that the equation admits an unique solution. The most important condition is the periodicity of the coefficients. In this sense, the model equation is the basis for our analysis. Nevertheless, the HM method introduced in section 3, is applicable to far more general cases.

The fundamental demand for our strategy to prove convergence of the later method, is the existence of a so called *two-scale homogenized equation* of the problem above. This is why we devote this subsection to recall the corresponding analytical results and its requirements.

To be sufficiently smooth for a subsequent numerical treatment, we assume from now on, that the coefficient functions are Lipschitz-continuous and that the initial value belongs to $H^1(\mathbb{R}^d)$. Moreover, these assumptions guarantee that we have a regular solution of the corresponding two-scale homogenized equation. This regularity will become important within later error estimates. The following assumptions are made:

Assumption 2.1 (General analytic assumptions)**.** We assume that the coefficient functions are Lipschitz-continuous, that the initial value is regular and that k is positive with average one, i.e.:

$$\begin{aligned} &A\in\left(H^{1,\infty}(0,\bar{T};H^{1,\infty}_\sharp(Y))\right)^{d\times d};\ \ A(t,y)\xi\cdot\xi\geq|\xi|^2\ \forall\xi\in\mathbb{R}^d \text{ a.e. in } (0,\bar{T})\times Y;\\ &b\in\left(H^{1,\infty}(0,\bar{T};H^{1,\infty}_\sharp(Y))\right)^{d};\ \ \nabla\cdot b(t,\cdot)=0 \text{ a.e. in } (0,\bar{T});\ \ v_0\in H^1(\mathbb{R}^d);\\ &k\in H^{1,\infty}(0,\bar{T};H^{1,\infty}_\sharp(Y));\ \ k>0;\ \ \int_Y k(t,y)\,dt = 1 \text{ everywhere in } [0,\bar{T}].\end{aligned}$$

Note that the regularity assumptions enable us to formulate the HMM with a pointwise evaluation of the coefficient functions. The condition that $k(t,\cdot)$ has average one, is just a normalization property, which simplifies the later results. The existence of k itself, may for instance be the result of a transformation from an advection-diffusion-reaction problem to an equation of the type above.

We introduce the following spaces:

Definition 2.2 (Analytic spaces)**.** For $0 \leq m < \infty$, $1 \leq p \leq \infty$ and for any $Y' = \prod_{i=0}^{d}(a_i, b_i) \subset \mathbb{R}^d$ with $a_i < b_i$, we define

$$\begin{aligned} C^0_\sharp(Y') &:= \{\phi \in C^0(Y') | \ \phi \ \text{is} \ Y' - \text{periodic}\}, \\ H^{m,p}_\sharp(Y') &:= \overline{C^0_\sharp(Y')}^{\|\cdot\|_{H^{m,p}(Y')}} \ \text{and} \\ \tilde{H}^1_\sharp(Y') &:= \{v \in H^1_\sharp(Y') \ | \fint_{Y'} v(y) \ dy = 0\}. \end{aligned}$$

For $Y = (0,1)^d$ we furthermore define the following Bochner-spaces:

$$\begin{aligned} I &:= H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d, H^1_\sharp(Y)), \\ I_0 &:= H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d, \tilde{H}^1_\sharp(Y)), \\ X^0(0,\bar{T}) &:= L^2(0,\bar{T}; H^1(\mathbb{R}^d)) \times L^2((0,\bar{T})\times\mathbb{R}^d, \tilde{H}^1_\sharp(Y)) \ \text{and} \\ X^1(0,\bar{T}) &:= H^1(0,\bar{T}; H^1(\mathbb{R}^d)) \times L^2((0,\bar{T})\times\mathbb{R}^d, \tilde{H}^1_\sharp(Y)). \end{aligned}$$

Let Ω be a domain, then $|\cdot|_{H^k(\Omega)}$ denotes the semi-norm on $H^k(\Omega)$ and the full norm is denoted by $\|\cdot\|_{H^k(\Omega)}$. Moreover, we introduce on $L^2(\Omega, H^k(Y))$:

$$|\Phi|_{L^2(\Omega,H^k(Y))} := \left(\int_\Omega |\Phi(x,\cdot)|^2_{H^k(Y)} \right)^{\frac{1}{2}} \text{ and } \|\Phi\|_{L^2(\Omega,H^k(Y))} := \sum_{l=0}^{k} |\Phi|_{L^2(\Omega,H^l(Y))}.$$

I_0 is a Hilbert space with respect to the norm $\|(\Phi,\phi)\|_{I_0} := |\Phi|_{H^1(\mathbb{R}^d)} + |\phi|_{L^2(\mathbb{R}^d,H^1(Y))}$ and $X^0(0,\bar{T})$ with $\|(\Phi,\phi)\|_{X^0(0,\bar{T})} := |\Phi|_{L^2(0,\bar{T};H^1(\mathbb{R}^d))} + |\phi|_{L^2((0,\bar{T})\times\mathbb{R}^d,H^1(Y))}$.

The following convergence was initially introduced by Marušić-Paloka and Piatnitski [23]:

Definition 2.3 (Two-scale convergence with drift)**.** Let $B \in H^1(0,\bar{T})^d$ be a given drift, $(u^\epsilon)_{\epsilon>0}$ a sequence in $L^2((0,\bar{T}) \times \mathbb{R}^d)$ and $u_0 \in L^2((0,\bar{T}) \times \mathbb{R}^d \times Y)$. Then we say u^ϵ is two-scale convergent with drift to u_0, if

$$\lim_{\epsilon\to 0} \int_0^{\bar{T}} \int_{\mathbb{R}^d} u^\epsilon(t,x)\Phi\left(t, x - \frac{B(t)}{\epsilon}, \frac{x}{\epsilon}\right) dx\, dt = \int_0^{\bar{T}} \int_{\mathbb{R}^d} \int_Y u_0(t,x,y)\Phi(t,x,y)\, dy\, dx\, dt$$

for all functions $\Phi \in L^2((0,\bar{T}) \times \mathbb{R}^d; C^0(Y))$.

In order to finally state the two-scale homogenized equation for problem (3), we still need some additional definitions. They will be used throughout the paper.

Definition 2.4. We define the space, and space-time averaged drift velocities $\bar{b}$ and B through

$$\bar{b}(t) := \fint_Y b(t,y)\, dy, \quad B(t) := \int_0^t \bar{b}(s)\, ds$$

and the parameter-depending bilinearform $E \in C^{0,1}([0,\bar{T}],\mathcal{L}(I,I'))$ by

$$
\begin{aligned}
E(t)\,((u_0,u_1),(\Phi_0,\phi_1)) :=\ & \int_{\mathbb{R}^d} \bar{b}(t)\cdot\nabla_x \Phi_0 \left(\int_Y k u_1\right) - \int_{\mathbb{R}^d}\int_Y \left(b(t,\cdot)\cdot\nabla_x\Phi_0\right) u_1 \\
- & \int_{\mathbb{R}^d} \bar{b}(t)\cdot\nabla_x u_0 \left(\int_Y k\phi_1\right) + \int_{\mathbb{R}^d}\int_Y b(t,\cdot)\cdot\left(\nabla_x u_0+\nabla_y u_1\right)\phi_1 \\
+ & \int_{\mathbb{R}^d}\int_Y A(t,\cdot)\left(\nabla_x u_0+\nabla_y u_1\right)\cdot\left(\nabla_x\Phi_0+\nabla_y\phi_1\right).
\end{aligned}
$$

Now we are prepared to formulate the main result. It is obtained by making the following asymptotic expansion ansatz for u^ϵ:

$$u^\epsilon(x) = u_0(t,x-\frac{B(t)}{\epsilon}) + \epsilon u_1\left(t,x-\frac{B(t)}{\epsilon},\frac{x}{\epsilon}\right) + \mathrm{O}(\epsilon^2).$$

In problem (3), we use test functions which are expected to be in resonance with the oscillations of u^ϵ, i.e. test functions of the form

$$\Phi^\epsilon(x) = \Phi_0(t,x-\frac{B(t)}{\epsilon}) + \epsilon\phi_1\left(t,x-\frac{B(t)}{\epsilon},\frac{x}{\epsilon}\right).$$

Forming the limit in the resulting weak formulation with $\epsilon\to 0$ yields the subsequent homogenization result (see [17] for details). The rather complicated structure of the two-scale operator E is a natural effect of the homogenization process. Indeed, decoupling the problem below and forming the effective macro problem, yields a parabolic equation, which only contains a diffusive part.

Theorem 2.5 (Two-scale homogenized equation with drift)**.** *Let Assumption 2.1 be fulfilled and let $(u^\epsilon)_{\epsilon>0}$ be the sequence of solutions of Problem (3). Then there exist functions $(u_0,u_1)\in X^1(0,\bar{T})$, such that we have the following convergence up to a subsequence:*

$$
\begin{aligned}
u^\epsilon &\to u_0 && \textit{two-scale with drift } B(t) \textit{ and}\\
\nabla u^\epsilon &\to \nabla_x u_0+\nabla_y u_1 && \textit{two-scale with drift } B(t).
\end{aligned}
$$

(u_0,u_1) is the unique solution of the homogenized problem

$$-\int_0^{\bar{T}} (u_0,\partial_t\Phi_0)_{L^2(\mathbb{R}^d)} + \int_0^{\bar{T}} E(t)((u_0,u_1),(\Phi_0,\phi_1)) = (v_0,\Phi_0(0,\cdot))_{L^2(\mathbb{R}^d)} \qquad (4)$$

for all $(\Phi_0,\phi_1)\in H^1(0,\bar{T};H^1(\mathbb{R}^d))\times L^2((0,\bar{T})\times\mathbb{R}^d,H^1_\sharp(Y))$, $\Phi_0(T,\cdot)=0$. Moreover, we have the following regularity for the solutions

$$
\begin{aligned}
&u_0\in H^1(0,\bar{T};H^1(\mathbb{R}^d))\cap L^2(0,\bar{T};H^2(\mathbb{R}^d)) \ \textit{ and}\\
&u_1\in L^2(0,\bar{T};H^1(\mathbb{R}^d,\tilde{H}^1_\sharp(Y)))\cap L^2((0,\bar{T})\times\mathbb{R}^d,H^2(Y))
\end{aligned}
$$

and the following estimate holds true

$$|u_1|_{L^2((0,\bar{T})\times\mathbb{R}^d,H^2(Y))} \le C|u_0|_{L^2(0,\bar{T};H^1(\mathbb{R}^d))} \le C\|v_0\|_{L^2(\mathbb{R}^d)}. \qquad (5)$$

Proof. A detailed proof can be found in [17], where the result is a combination of Theorem 3.1, Proposition 3.3 and Remark 3.6. □

2.2. **The discrete setting.** Before we formulate the heterogeneous multiscale method for our types of problems, we still need to make several definitions in order to introduce the discrete spaces that we are dealing with.

Therefore, let $\mathcal{T}_H = \{T_j | j \in J\}$ be a regular simplicial partition of $\mathbb{R}^d$, and $\{(q_i, x_i) | i \in Q_j\}$ a given quadrature rule on $T_j \in \mathcal{T}_H$ with weights q_i and quadrature points x_i. By $\mathcal{T}_h$ we denote a regular periodic partition of Y with index set K such that $\mathcal{T}_h = \{S_k | k \in K\}$. Furthermore we define the $\frac{1}{2}$-tuple in $\mathbb{R}^d$ by $v_{\frac{1}{2}} = (\frac{1}{2}, ..., \frac{1}{2})$ and the ϵ-scaled unit-cells, centered around a quadrature point x_i by $Y_{i,\epsilon} := \{x_i - \epsilon v_{\frac{1}{2}} + \epsilon y \ | y \in Y\}$. The associated bijection $x_i^\epsilon : Y \to Y_{i,\epsilon}$ is given by $x_i^\epsilon(y) := x_i - \epsilon v_{\frac{1}{2}} + \epsilon y, \ \forall y \in Y$ and $y_i^\epsilon : Y_{i,\epsilon} \to Y$ the corresponding inverse mapping defined by $y_i^\epsilon(x) := \frac{x - x_i + \epsilon v_{\frac{1}{2}}}{\epsilon}, \ \forall x \in Y_{i,\epsilon}$. For simplification we will use equidistant time steps. Therefore, we define $t^n := n\triangle t$, where $\triangle t := \frac{\bar{T}}{N}$ denotes the step size and $N \in \mathbb{N}$ the maximum number of time steps. Moreover, we introduce the following spaces:

Definition 2.6 (Discrete spaces)**.** For the multiscale method we define

$$\begin{aligned}
V_H^l := V_H^l(\mathbb{R}^d) \quad &:= \quad \{\Phi_H \in H^1(\mathbb{R}^d) \cap C^0(\mathbb{R}^d) \ | \Phi_{H_{|T}} \in \mathbb{P}^l(T) \ \forall T \in \mathcal{T}_H\}; \\
W_h^m(Y) \quad &:= \quad \{\phi_h \in \tilde{H}_\sharp^1(Y) \cap C^0(Y) \ | \ \phi_{h|S} \in \mathbb{P}^m(S) \ \forall S \in \mathcal{T}_h\}; \\
W_h^m(Y_{i,\epsilon}) \quad &:= \quad \{\phi_h \in \tilde{H}_\sharp^1(Y_{i,\epsilon}) \cap C^0(Y_{i,\epsilon}) | \ (\phi_h \circ x_i^\epsilon)_{|S} \in \mathbb{P}^m(S) \ \forall S \in \mathcal{T}_h\}; \\
V_H^l(\mathbb{R}^d, W_h^m(Y)) \quad &:= \quad \{\phi_h \in L^2(\mathbb{R}^d, \tilde{H}_\sharp^1(Y)) \ | \phi_h(\cdot, y)_{|T} \in \mathbb{P}^l(T) \ \forall T \in \mathcal{T}_H, y \in Y; \\
& \qquad \phi_h(x, \cdot) \in W_h^m(Y) \ \forall x \in \mathbb{R}^d\};
\end{aligned}$$

and for a reformulation of the method we also introduce

$$\begin{aligned}
\tilde{I}_H &:= V_H^1(\mathbb{R}^d) \times V_H^0(\mathbb{R}^d, W_h^1(Y)), \text{ the space of solutions per time step;} \\
I_H &:= V_H^1(\mathbb{R}^d) \times L^2(\mathbb{R}^d, W_h^1(Y)), \text{ the space of test functions per time step;} \\
\tilde{X}_H(0,\bar{T}) &:= V_{\triangle t}^0(0,\bar{T}; V_H^1(\mathbb{R}^d)) \times V_{\triangle t,H}^0((0,\bar{T}) \times \mathbb{R}^d, W_h^1(Y)), \text{ the solution space;} \\
X_H(0,\bar{T}) &:= V_{\triangle t}^0(0,\bar{T}; V_H^1(\mathbb{R}^d)) \times L^2((0,\bar{T}) \times \mathbb{R}^d, W_h^1(Y)), \text{ the test space.}
\end{aligned}$$

Here $V_{\triangle t}^0(0,\bar{T})$ denotes the space of piecewise constant functions on every interval $[n\triangle t, (n+1)\triangle t] \subset [0, \bar{T}]$. $V_{\triangle t,H}^0((0,\bar{T}) \times \mathbb{R}^d)$ is the space of piecewise constant functions on elements $[n\triangle t, (n+1)\triangle t] \times T$, $T \in \mathcal{T}_H$. $\tilde{X}_H$ and X_H should be seen as different approximations of the original solution space X^0. After reformulation, $\tilde{X}_H$ is the solution space for our HMM-approximation. In the following section, we will see that the reformulation of the HMM yields the possibility to use test functions, which partially contain a very general L^2-part, therefore we also introduce X_H. This fact is used for achieving our a-priori error estimate.

3. **Heterogeneous multiscale method for advection-diffusion problems.** We are now prepared to derive a suitable multiscale finite element method for a general (possibly non-periodic) problem of the following kind:
find $u^\epsilon \in H^1(0, \bar{T}; H^1(\mathbb{R}^d))$ with

$$\begin{aligned}
&\int_0^{\bar{T}} \int_{\mathbb{R}^d} k\left(t, \frac{x}{\epsilon}\right) \partial_t u^\epsilon(t,x) \Phi(t,x) + A^\epsilon(t,x) \nabla u^\epsilon(t,x) \cdot \nabla \Phi(t,x) \, dx \, dt \qquad (6) \\
&\quad + \int_0^{\bar{T}} \int_{\mathbb{R}^d} \epsilon^{-1} b^\epsilon(t,x) \cdot \nabla u^\epsilon(t,x) \Phi(t,x) \, dx \, dt = 0 \quad \forall \Phi \in H^1(0, \bar{T}, H^1(\mathbb{R}^d))
\end{aligned}$$

and $u^\epsilon(0,\cdot) = v_0$. We demand that b^ϵ is divergence-free. If there is additionally some cell size $1 \gg \delta > 0$ with $\int_{x+[-\delta,\delta]^d} b^\epsilon(t,y)\,dy = 0$ for all $(t,x)\in[0,\bar{T}]\times\mathbb{R}^d$ we do not need further assumptions on A^ϵ. If b^ϵ does not have local zero average, we assume that A^ϵ and b^ϵ are only micro-scale functions, i.e. they only show a microscopic behaviour and are constant on the macro-scale (for fixed t). However, these restrictions are not necessarily needed. At the moment they should be regarded as a simplification for the method. If there is absolutely no restriction on the scale separated functions b^ϵ and A^ϵ (except div $b^\epsilon = 0$), the coefficients need to be premodified according to the macroscopic drift. This may for instance involve a scale separation by means of multiresolution analysis.

Moreover, we note that the specific structure of $k^\epsilon(t,x) = k(t, \frac{x}{\epsilon})$ with average $\int_Y k(t,y)\ dy = 1$, is also not a real restriction, but a simplification. The case with a completely general k^ϵ yields no further difficulties but does not make sense in the formulation above. This is due to the fact that the coefficient function k^ϵ only occurs when the considered problem is a transformation of an originally more general problem with reaction. This transformation however, is only possible under certain conditions. Therefore, if the other coefficient functions fulfill these conditions, we always have that k^ϵ is of the form $k(\cdot, \frac{\cdot}{\epsilon})$.

3.1. **Motivation for the formulation of the multiscale method.**

In this subsection, we a heuristic approach of how to formulate the multiscale method. It should not be regarded as a proof, but only as a motivation for the scheme that we find in Definition 3.1.

Let $\Phi \in H^1((0,\bar{T}) \times \mathbb{R}^d)$ be a test function, then we start with the variational formulation of (6):

$$\int_{\mathbb{R}^d}\int_0^{\bar{T}} k^\epsilon \partial_t u^\epsilon \Phi + \int_{\mathbb{R}^d}\int_0^{\bar{T}} A^\epsilon \nabla u^\epsilon \cdot \nabla\Phi + \frac{1}{\epsilon}\int_{\mathbb{R}^d}\int_0^{\bar{T}} (b^\epsilon \cdot \nabla u^\epsilon)\,\Phi = 0$$

and u^ϵ fulfilling the initial-boundary condition. Using that b^ϵ is divergence-free we obtain:

$$\int_{\mathbb{R}^d}\int_0^{\bar{T}} k^\epsilon \partial_t u^\epsilon \Phi + \int_{\mathbb{R}^d}\int_0^{\bar{T}} A^\epsilon \nabla u^\epsilon \cdot \nabla\Phi - \frac{1}{\epsilon}\int_{\mathbb{R}^d}\int_0^{\bar{T}} (b^\epsilon \cdot \nabla \Phi)\,u^\epsilon = 0.$$

We make a finite-element approach with quadrature formula for this problem. Keeping in mind that u^ϵ and the coefficient functions contain fine-scale oscillations, we naturally formulate the following equation:

$$(u_H^{n+1}, \Phi_H)_{L^2(\mathbb{R}^N)} \quad = \quad (u_H^n, \Phi_H)_{L^2(\mathbb{R}^N)} + \triangle t A_H^{n+1}(u_H^{n+1}, \Phi_H) \qquad \forall \Phi_H \in V_H^1$$

with

$$\begin{aligned} A_H^n(u_H, \Phi_H) := \quad & \sum_{j\in J}\sum_{i\in Q_j} q_i \fint_{Y_{i,\epsilon}} A_h^\epsilon(t^n, x)\nabla_x R_i^{(n)}(u_H)(x)\cdot\nabla_x \Phi_H(x_i)\ dx \\ - \quad & \sum_{j\in J}\sum_{i\in Q_j} q_i \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^\epsilon(t^n, x)\cdot\nabla_x \Phi_H(x_i)\ \ R_i^{(n)}(u_H)(x)\ dx. \end{aligned}$$

Here, A_h^ϵ and b_h^ϵ denote adequate approximations of A^ϵ and b^ϵ (see Definition 3.1 for details). $R_i^{(n)}$ denotes a reconstruction operator, such that $R_i^{(n)}(u_H)$ approximates

u^ϵ. Note that k does not occur in the scalar products of the kind $(u_H^n, \Phi_H)_{L^2(\mathbb{R}^N)}$, since it has average 1 on the micro-scale:

$$(u_H^n, \Phi_H)_{L^2(\mathbb{R}^N)} \approx \sum_{j\in J}\sum_{i\in Q_j} q_i u_H^n(x_i)\Phi_H(x) = \sum_{j\in J}\sum_{i\in Q_j} q_i \fint_{Y_{i,\epsilon}} k^\epsilon u_H^n(x_i)\Phi_H(x).$$

There are two questions that arise with this approach:

1. Why do we approximate $-\int_{\mathbb{R}^d}\int_0^T (b^\epsilon\cdot\nabla\Phi)\,u^\epsilon$ instead of $\int_{\mathbb{R}^d}\int_0^T (b^\epsilon\cdot\nabla u^\epsilon)\,\Phi$? Does it make a difference?
2. How do we determine the reconstruction operator $R_i^{(n)}$?

The answer to the first question is very much related to the answer of the second question. In general, the reconstructions in a heterogeneous multiscale method have difficulties to capture terms of order $O(\epsilon)$. For instance: Assume that we make the ansatz $u^\epsilon(x) = u_0(x)+\epsilon u_1(x,\frac{x}{\epsilon})$, then $\nabla R_i(u_H)$ approximates $\nabla_x u_0+\nabla_y u_1$ instead of $\nabla_x u_0+\epsilon\nabla_x u_1+\nabla_y u_1$. Normally, this is not problematic, since ϵ tends to zero. But in a term that is scaled with $\frac{1}{\epsilon}$, it becomes significant. On the other hand, in the reconstruction $R_i(u_H)$ itself (an approximation of $u_0+\epsilon u_1$) the term of order ϵ is still existent. Therefore, $-\sum_{j\in J}\sum_{i\in Q_j} q_i\fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^\epsilon(t^n,x)\cdot\nabla_x\Phi_H(x_i)\;R_i^{(n)}(u_H)(x)\;dx$ is expected to be a relatively exact approximation (it also captures the $O(\epsilon)$-terms), whereas $\sum_{j\in J}\sum_{i\in Q_j} q_i\fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^\epsilon(t^n,x)\cdot\nabla_x R_i^{(n)}(u_H)(x)\;\Phi_H(x)\;dx$ does not approximate the right term (it does not capture the $O(\epsilon)$-terms).

Now we focus on question 2. The discrete problem for the determination of the reconstructions is given by a discretization of the local resonance condition, namely find $R_i^{(n)}(u_H) \in u_H + W_h^m(Y_{i,\delta})$ with

$$\begin{aligned}&\int_{Y_{i,\delta}} A_h^\epsilon(t^n,x)\nabla_x R_i^{(n)}(u_H)(x)\cdot\nabla_x\phi_h(x)\;dx\\ &\quad+\int_{Y_{i,\delta}}\frac{1}{\epsilon}b_h^\epsilon(t^n,x)\cdot\nabla_x R_i^{(n)}(u_H)(x)\;\phi_h(x)\;dx = 0, \qquad \forall\phi_h\in W_h^1(Y_{i,\delta}).\end{aligned}$$

This equation can be interpreted as that the micro-scale oscillations of the reconstruction $R_i(u_H)$ are in resonance with all functions $\phi_h \in W_h^1(Y_{i,\delta})$. This is what we expect if we make the ansatz $u^\epsilon(t,x) = u_0(t,x) + \epsilon u_1(t,x,\frac{x}{\epsilon}) + O(\epsilon^2)$. Note that the test function ϕ_h should be interpreted as a shifted test function scaled with ϵ, i.e. it is of the type $\epsilon\tilde{\phi}_h$. Therefore, the multiplication with $\frac{1}{\epsilon}$ is neutralized and we do not deal with the same problem as in the global equation. This means that using $\int_{Y_{i,\delta}}\frac{1}{\epsilon}b_h^\epsilon(t^n,x)\cdot\nabla_x R_i^{(n)}(u_H)(x)\;\phi_h(x)\;dx$ instead of $-\int_{Y_{i,\delta}}\frac{1}{\epsilon}b_h^\epsilon(t^n,x)\cdot\nabla\phi_h(x)\;R_i^{(n)}(u_H)(x)\;dx$ makes sense in this formulation. Both expressions only differ in a term of order $O(\epsilon)$.

Using the derived method as described above will still produce wrong approximations for u^ϵ. Why? Again we need to focus on the macro-scale part with the $\frac{1}{\epsilon}$-dependence. Defining the *local centered reconstruction* by:

$$\overline{R_i^{(n)}}(u_H)(x) := R_i^{(n)}(u_H)(x) - (u_H(x) - u_H(x_i)),$$

we look at the difference between the expressions:

$$\fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; R_i^{(n)}(u_H)(x) \;\; dx \;\; \text{and} \tag{7}$$

$$\fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; \overline{R_i^{(n)}}(u_H)(x) \;\; dx. \tag{8}$$

Even if they seem almost identical, the difference may be crucial. Both terms differ in the separation of the scales. In (7) the scales seem to be 'less separated' than in (8). For instance, in (8) we are essentially dealing with an average over the micro-scale behaviour. In (7) on the other hand, the macro-scale behaviour has an influence of order ϵ on the average. At first view, an $\mathrm{O}(\epsilon)$-discrepancy seems to be negligible, but again the $\frac{1}{\epsilon}$-scaling can produce a significant difference. The following argumentation is to emphasize this. Let us assume that $b^{\epsilon}(t,x) = b(t, \frac{x}{\epsilon})$ with $b(t,\cdot)$ being Y-periodic and having zero average, then we have for a suitable approximation b_h:

$$\fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; u_H(x_i)\, dx = u_H(x_i) \;\; \nabla_x \Phi_H(x_i) \cdot \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x)\, dx = 0. \tag{9}$$

If we furthermore define the fine-scale part by $\mathcal{K}_i^{(n)}(u_H) := R_i^{(n)}(u_H) - u_H$, we have:

$$\begin{aligned}(8) &= \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; \Big(u_H(x_i) + \mathcal{K}_i^{(n)}(u_H)\Big) \; dx \\ &\overset{(9)}{=} \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; \mathcal{K}_i^{(n)}(u_H)\, dx.\end{aligned}$$

On the other hand for (7):

$$\begin{aligned}(7) &= \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; \Big(u_H(x) + \mathcal{K}_i^{(n)}(u_H)\Big) \;\; dx \\ &\overset{(9)}{=} \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; \Big(u_H(x) - u_H(x_i) + \mathcal{K}_i^{(n)}(u_H)\Big) \;\; dx \\ &= \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; (u_H(x) - u_H(x_i)) \;\; dx \\ &\quad + \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; \mathcal{K}_i^{(n)}(u_H) \;\; dx.\end{aligned}$$

Thus, (7) and (8) differ in:

$$\begin{aligned}&\fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; (u_H(x) - u_H(x_i)) \;\; dx \\ &\quad = \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} b_h^{\epsilon}(t^n,x) \cdot \nabla_x \Phi_H(x_i) \;\; (\nabla u_H(x_i) \cdot (x - x_i)) \;\; dx.\end{aligned} \tag{10}$$

Since we only have $|x - x_i| \leq \epsilon$ and at the same time a scaling with $\frac{1}{\epsilon}$, the difference (10) is neither equal to zero nor does it converge to zero. Therefore, (7) and (8) are obviously different. In the following we work with (8), since a clearer separation of the scales corresponds more with our approach for u^{ϵ}. As we will see later on, this choice is the 'right' choice.

We simplify the subsequent considerations by assuming that the macroscopic drift B is only time dependent. Since we expect the solution u^{ϵ} to have a large drift, we integrate this presumption into the method itself. Just like in the analytical contemplations, we make the ansatz:

$$u^\epsilon(t,x) = u_0\Big(t, x - \frac{B(t)}{\epsilon}\Big) + \epsilon u_1\Big(t, x - \frac{B(t)}{\epsilon}, \big(x, \epsilon^{-1}\big)\Big) + O(\epsilon^2). \tag{11}$$

Here, the last component of u_1 describes the microscopic behaviour of u^ϵ (without being necessarily periodic).

Instead of approximating $U_0(t,x) = u_0\Big(t, x - \frac{B(t)}{\epsilon}\Big)$, the multiscale method will be designed to approximate only $u_0(t,x)$ by using an approach of characteristics. With this approach we have the advantage that the strongly dominating part of order $\frac{1}{\epsilon}$ can be erased in our discrete problems. Moreover, we do not have to use small time step sizes to capture the drift.

In order to incorporate these ideas into the method, we proceed heuristically. Equation (11) suggests to test with
$\Phi^\epsilon(t,x) = \Phi^0\Big(t, x - \frac{B(t)}{\epsilon}\Big) + \epsilon\phi^1\Big(t, x - \frac{B(t)}{\epsilon}, \big(x, \epsilon^{-1}\big)\Big)$ to determine the changes on the macro-scale and on the micro-scale. Terms which are of the order $O(\epsilon)$ are neglected.

$$\begin{aligned} \partial_t \Phi^\epsilon(t,x) = \;& \partial_t \Phi^0\Big(t, x - \frac{B(t)}{\epsilon}\Big) - \frac{1}{\epsilon}\bar{b}(t)\cdot\nabla_x \Phi^0\Big(t, x - \frac{B(t)}{\epsilon}\Big) \\ & - \frac{1}{\epsilon}\bar{b}(t)\cdot\epsilon\nabla_x \phi^1\Big(t, x - \frac{B(t)}{\epsilon}, \big(x, \epsilon^{-1}\big)\Big) + O(\epsilon). \end{aligned} \tag{12}$$

Note that the average $\bar{b}(t)$ of the advective part b^ϵ is equal to the derivative of the macroscopic drift $B(t)$. k^ϵ shall be disregarded for the moment, since it produces no crucial changes. Since

$$\int_{\mathbb{R}^d}\int_0^{\bar{T}} \partial_t u^\epsilon \Phi^\epsilon \, dt\, dx \quad = \quad -\int_{\mathbb{R}^d}\int_0^{\bar{T}} u^\epsilon \partial_t \Phi^\epsilon \, dt\, dx$$

we conclude with (12) that there should be an additional term on the macro-scale, that behaves like $(u_0 + \epsilon u_1)\frac{1}{\epsilon}\bar{b}\nabla_x \Phi^0$. We therefore need to add the following part to our method

$$\sum_{j\in J}\sum_{i\in Q_j} q_i \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon}\bar{b}(t^n)\cdot\nabla_x \Phi_H(x_i)\ \overline{R_i^{(n)}}(u_H)(x)\ dx.$$

On the micro-scale we observe that there should be an additional term behaving like $\int_{\mathbb{R}^d}\int_Y u_0 \nabla_x \phi^1$. Since $\int_{\mathbb{R}^d}\int_Y u_0\nabla_x\phi^1 = -\int_{\mathbb{R}^d}\int_Y \nabla_x u_0 \phi^1$, we add
$-\int_{Y_{i,\delta}} \frac{1}{\epsilon}\bar{b}(t^n)\cdot\nabla_x u_H(x)\ \phi_h(x)\ dx$ to the micro-scale equation. This concludes our considerations and we are ready to formulate our multiscale finite element method.

3.2. **Formulation of the HMM for the general non-periodic case.**

In this subsection, we state the heterogeneous multiscale finite element method for advection-diffusion problems. No periodicity is assumed for this part. The HMM reads as follows:

Definition 3.1 (HMM for advection-diffusion problems with large drift)**.** Assume that b^ϵ is a divergence-free advection velocity, then we define the HMM approximation U_H of u^ϵ by

$$U_H(t^n, x_i) := u_H^n(x_i - \frac{1}{\epsilon}\int_0^{t^n}\fint_{Y_{i,\delta}} b_h^\epsilon(y,s)\ dy\ ds)$$

where $u_H^{n+1} \in V_H^l$ is defined as the solution of

$$(u_H^{n+1}, \Phi_H)_{L^2(\mathbb{R}^N)} + \triangle t A_H^{n+1}(u_H^{n+1}, \Phi_H) = (u_H^n, \Phi_H)_{L^2(\mathbb{R}^N)} \qquad \forall \Phi_H \in V_H^l,$$

with

$$A_H^n(u_H, \Phi_H) := \sum_{j\in J} \sum_{i \in Q_j} q_i \fint_{Y_{i,\epsilon}} A_h^\epsilon(t^n, x) \nabla_x R_i^{(n)}(u_H)(x) \cdot \nabla_x \Phi_H(x_i) \ dx$$
$$+ \sum_{j\in J} \sum_{i \in Q_j} q_i \fint_{Y_{i,\epsilon}} \frac{1}{\epsilon} \left(k_h^\epsilon(t^n, x) \bar{b}_h^{\epsilon,i}(t^n) - b_h^\epsilon(t^n, x) \right) \cdot \nabla_x \Phi_H(x_i) \ \overline{R_i^{(n)}}(u_H)(x) \ dx.$$

Here, A_h^ϵ, b_h^ϵ and k_h^ϵ are assumed to be suitable approximations of A^ϵ, b^ϵ and k^ϵ. If these coefficient functions are sufficiently regular, we may for instance use $A_h^\epsilon(t, \cdot)_{|x_i^\epsilon(S_k)} := A^\epsilon(t, x_i^\epsilon(y_k))$ and $b_h^\epsilon(t, \cdot)_{|x_i^\epsilon(S_k)} := b^\epsilon(t, x_i^\epsilon(y_k))$, where y_k denotes the barycenter of S_k. Moreover, we define $\bar{b}_h^{\delta,i}(t) := \fint_{Y_{i,\delta}} b_h^\epsilon(t, y) \ dy$.

The *local centered reconstructions* are given by:

$$\overline{R_i^{(n)}}(\Phi_H)(x) := R_i^{(n)}(\Phi_H)(x) - (\Phi_H(x) - \Phi_H(x_i)),$$

and for $\Phi_H \in V_H^l$ the *local reconstruction operator* $R_i^{(n)}$ itself is defined by the solutions $R_i^{(n)}(\Phi_H) \in \Phi_H + W_h^m(Y_{i,\delta})$ of

$$\int_{Y_{i,\delta}} A_h^\epsilon(t^n, x) \nabla_x R_i^{(n)}(\Phi_H)(x) \cdot \nabla_x \phi_h(x) \, dx + \int_{Y_{i,\delta}} \frac{1}{\epsilon} b_h^\epsilon(t^n, x) \cdot \nabla_x R_i^{(n)}(\Phi_H)(x) \ \phi_h(x) \, dx$$
$$= \int_{Y_{i,\delta}} k_h^\epsilon(t^n, x) \ \frac{1}{\epsilon} \bar{b}_h^{\delta,i}(t^n) \cdot \nabla_x \Phi_H(x) \ \phi_h(x) \, dx, \qquad \forall \phi_h \in W_h^m(Y_{i,\delta}).$$

The initial value $u_H^0 = v_H^0$ is given by a suitable discretization of v_0. For the parameter δ we furthermore assume $\delta \geq \epsilon$. An expedient choice for the periodic case could be $\delta = \epsilon$, for the non-periodic case $\delta = m\epsilon, m > 1$.

In Definition 3.1 we assume that $\fint_{Y_{i,\delta}} k_h^\epsilon(t^n, x) \ dx = 1$ for all $Y_{i,\delta}$ and for all time steps t^n. If this is not the case, the HMM needs to be modified according to a new drift of the form

$$B_h(t^n) = \int_0^{t^n} \frac{\fint_{Y_{i,\delta}} b_h^\epsilon(s, x) \, dx}{\fint_{Y_{i,\delta}} k_h^\epsilon(s, x) \, dx} ds.$$

This can be done in a straightforward way.

3.3. **The HMM and its reformulation for the periodic case.** In section 4, we derive an a-priori error estimate for the HMM defined in 3.1. To do so, we need to restrict ourselves to the case of periodic coefficient functions. In this subsection, we therefore introduce a simplified formulation of the HMM in the periodic setting. We show that it is equivalent to a direct discretization of the homogenized problem (4). This result yields the basis for the analysis in section 4.

In Definition 3.2 below, we only use a Newton-Cotes quadrature formula of order zero. Note that this is not a real constraint. For the case of quadrature formulas of higher order, additional error terms occur, which depend on this order. Reformulations of the HMM will therefore contain an approximation error related to the quadrature. The following method is merely a simplification of the method in Definition 3.1. In the following, we will always refer to this simplification.

Definition 3.2. Let x_j be defined as the barycenter of the macro-grid element $T_j \in \mathcal{T}_H$ and y_k the barycenter of micro-grid element $S_k \in \mathcal{T}_h$. Under Assumption 2.1, we furthermore define the discrete approximations of A^ϵ, b^ϵ and k^ϵ by

$$A_h^\epsilon(t,x) := A\left(t^n, \frac{x_j^\epsilon(y_k)}{\epsilon}\right),\ b_h^\epsilon(t,x) := b\left(t^n, \frac{x_j^\epsilon(y_k)}{\epsilon}\right),\ k_h^\epsilon(t,x) := k\left(t^n, \frac{x_j^\epsilon(y_k)}{\epsilon}\right)$$

for $(t,x) \in [t^n, t^{n+1}) \times x_j^\epsilon(S_k)$. Moreover, we make use of $A_h(t,y) := A_h^\epsilon(t, \epsilon y)$. b_h and k_h are defined in analogy. Since b_h is a Y-periodic function we can simplify the definition of $\bar{b}_h^{\delta,j}$ to:

$$\bar{b}_h^{\delta,j}(t) := \bar{b}_h(t) = \fint_Y b_h(t,y)\ dy.$$

In the periodic setting, Definition 3.1 can be expressed as follows:

Definition 3.3 (HMM for periodic coefficient functions)**.** In the case of periodic coefficient functions we will use the following version of the HMM. Here the HMM approximation U_H of u^ϵ is given by

$$U_H(t^n, x) := u_H^n(x - \frac{1}{\epsilon}\int_0^{t^n} \fint_Y b_h(y,s)\ dy\ ds),$$

where $u_H^{n+1} \in V_H^1$ is defined as the solution of:

$$(u_H^n, \Phi_H)_{L^2(\mathbb{R}^N)} \quad = \quad (u_H^{n+1}, \Phi_H)_{L^2(\mathbb{R}^N)} + \triangle t A_H^{n+1}(u_H^{n+1}, \Phi_H) \qquad \forall \Phi_H \in V_H^1,$$

with

$$\begin{aligned} A_H^n(u_H, \Phi_H) &:= \sum_{j\in J} |T_j| \fint_{Y_{j,\epsilon}} A_h^\epsilon(t^n,x) \nabla_x R_j^{(n)}(u_H)(x) \cdot \nabla_x \Phi_H(x_j)\ dx \\ &+ \sum_{j\in J} |T_j| \fint_{Y_{j,\epsilon}} \frac{1}{\epsilon}\left(k_h^\epsilon(t^n,x)\bar{b}_h(t^n) - b_h^\epsilon(t^n,x)\right) \cdot \nabla_x \Phi_H(x_j)\ \overline{R_j^{(n)}}(u_H)(x)\ dx. \end{aligned}$$

x_j, A_h^ϵ, b_h^ϵ, k_h^ϵ and $\bar{b}_h$ are given by Definition 3.2. The local centered reconstructions are defined by:

$$\overline{R_j^{(n)}}(\Phi_H)(x) := R_j^{(n)}(\Phi_H)(x) - (\Phi_H(x) - \Phi_H(x_j)),$$

and for $\Phi_H \in V_H^1$ the local reconstruction operator $R_j^{(n)}$ itself is defined by the solutions $R_j^{(n)}(\Phi_H) \in \Phi_H + W_h^1(Y_{j,\epsilon})$ of

$$\begin{aligned} &\int_{Y_{j,\epsilon}} A_h^\epsilon(t^n,x) \nabla_x R_j^{(n)}(\Phi_H)(x) \cdot \nabla_x \phi_h(x)\, dx + \int_{Y_{j,\epsilon}} \frac{1}{\epsilon} b_h^\epsilon(t^n,x) \cdot \nabla_x R_j^{(n)}(\Phi_H)(x)\, \phi_h(x)\, dx \\ &\quad = \int_{Y_{j,\epsilon}} k_h^\epsilon(t^n,x)\ \frac{1}{\epsilon}\bar{b}_h(t^n) \cdot \nabla_x \Phi_H(x)\ \phi_h(x)\ dx, \qquad \forall \phi_h \in W_h^1(Y_{j,\epsilon}). \end{aligned}$$

The initial value u_H^0 is given by $u_H^0 := v_H^0 := \mathcal{I}_H(v^0)$.

To prepare for the numerical analysis, we now draw our attention to a reformulation of this method. For this purpose, we introduce the bilinear forms E_H and G^N:

Definition 3.4. We define the operator $E_H^n \in \mathcal{L}(\tilde{I}_H, I'_H)$ by:

$$\begin{aligned} E_H^n((u_H,u_h),(\Phi_H,\phi_h)) &:= \int_{\mathbb{R}^d}\int_Y b_h(t^n,y)\cdot(\nabla_x u_H(x)+\nabla_y u_h(x,y))\ \ \phi_h(x,y)\,dy\,dx \\ &+ \int_{\mathbb{R}^d}\int_Y \left(k_h(t^n,y)\bar{b}_h(t^n)-b_h(t^n,y)\right)\cdot\nabla_x\Phi_H(x)\ \ u_h(x,y)\,dy\,dx \\ &- \int_{\mathbb{R}^d}\int_Y k_h(t^n,y)\ \ \bar{b}_h(t^n)\cdot\nabla_x u_H(x)\ \ \phi_h(x,y)\,dy\,dx \\ &+ \int_{\mathbb{R}^d}\int_Y A_h(t^n,y)\left(\nabla_x u_H(x)+\nabla_y u_h(x,y)\right)\cdot\left(\nabla_x\Phi_H(x)+\nabla_y\phi_h(x,y)\right)\,dy\,dx \end{aligned}$$

and the parameter dependent bilinear form $E_H \in V^0_{\triangle t}([0,\bar{T}],\mathcal{L}(\tilde{I}_H,I'_H))$ by

$$E_H(t)\,((u_H,u_h),(\Phi_H,\phi_h)) := E_H^{n+1}\,((u_H,u_h),(\Phi_H,\phi_h)) \ \text{ for } \ t\in(t^n,t^{n+1}].$$

Moreover, we define the jumps over t^n by:

$$[u]_n := u_+^n - u_-^n, \ \text{ where } \ u_+^n := \lim_{t\searrow t^n} u(t,\cdot), \ \ u_-^n := \lim_{t\nearrow t^n} u(t,\cdot)$$

and a corresponding space by:

$$X_{\triangle t} := \{(\Phi_0,\phi_1)\in X^0(0,\bar{T})\ \ |\Phi_{0|[t^n,t^{n+1}]}\in\mathbb{P}^0(t^n,t^{n+1};H^1(\mathbb{R}^d))\}. \tag{13}$$

For simplification we furthermore denote for $n\geq 0$ and $(\Phi_H,\phi_h)\in X_H(0,\bar{T})$

$$\Phi_H^{n+1} := (\Phi_H)_+^n = (\Phi_H)_-^{n+1}. \tag{14}$$

The bilinear form $G^N : X^1(0,\bar{T})\times X_{\triangle t}\to\mathbb{R}$ is given by:

$$\begin{aligned} G^N((u_0,u_1),(\Phi_0,\phi_1)) &:= \sum_{n=1}^{N-1}([u_0]_n,(\Phi_0)_+^n)_{L^2(\mathbb{R}^d)} + ((u_0)_+^0,(\Phi_0)_+^0)_{L^2(\mathbb{R}^d)} \\ &+ \sum_{n=0}^{N-1}\left(\int_{t^n}^{t^{n+1}}(\partial_t u_0,\Phi_0)_{L^2(\mathbb{R}^d)} + E(t)((u_0,u_1),(\Phi_0,\phi_1))\right) \end{aligned}$$

and analogously for the discrete case $G_H^N : \tilde{X}_H(0,\bar{T})\times X_H(0,\bar{T})\to\mathbb{R}$:

$$\begin{aligned} G_H^N((u_H,u_h),(\Phi_H,\phi_h)) &:= \sum_{n=1}^{N-1}([u_H]_n,(\Phi_H)_+^n)_{L^2(\mathbb{R}^d)} + ((u_H)_+^0,(\Phi_H)_+^0)_{L^2(\mathbb{R}^d)} \\ &+ \sum_{n=0}^{N-1}\left(\int_{t^n}^{t^{n+1}}(\partial_t u_H,\Phi_H)_{L^2(\mathbb{R}^d)} + E_H(t)((u_H,u_h),(\Phi_H,\phi_h))\right). \end{aligned}$$

The following theorem shows that in the periodic case, the HMM is equivalent to a discretization of the two-scale equation (4) by means of a Discontinuous Galerkin Time Stepping Method with quadrature. In this spirit, u_H is an approximation of the macro-scale portion u_0, whereas $\mathcal{K}_h(u_H)$ (defined in the subsequent theorem) approximates the micro-scale part u_1. This fact will help us to derive a corresponding a-priori error estimate.

Theorem 3.5 (Reformulation of HMM). *Suppose that $H \gg \epsilon$ and let U_H, u_H^n and $R_j^{(n)}(u_H^n)$ be given by Definition 3.3. We furthermore define $(u_H,\mathcal{K}_h(u_H))\in\tilde{X}_H(0,\bar{T})$ by $u_{H|(t^n,t^{n+1}]} := u_H^{n+1}$ and $\mathcal{K}_h(u_H)_{|(t^n,t^{n+1}]} := \mathcal{K}_h^{(n+1)}(u_H^{n+1})$, where $\mathcal{K}_h^{(n)}(u_H^n)\in V_H^0(\mathbb{R}^d,W_h^1(Y))$ is given by $\mathcal{K}_h^{(n)}(u_H^n)(x,y)_{|T_j\times Y} := \frac{1}{\epsilon}\mathcal{K}_j^{(n)}(u_H^n)(\epsilon y)$ and*

$\mathcal{K}_j^{(n)}(u_H^n) \in W_h^1(Y_{j,\epsilon})$ *by* $\mathcal{K}_j^{(n)}(u_H^n) := R_j^{(n)}(u_H^n) - u_H^n$. *Note that any periodic function should be seen as its extension to the whole* $\mathbb{R}^d$*, so that the preceding definitions make sense. With these assumptions, we have that* $(u_H, \mathcal{K}_h(u_H)) \in \tilde{X}_H(0, \bar{T})$ *is a solution of*

$$G_H^N((u_H, \mathcal{K}_h(u_H)), (\Phi_H, \phi_h)) = (v_H^0, (\Phi_H)_+^0)$$

for all $(\Phi_H, \phi_h) \in X_H(0, \bar{T})$ *and for all* N*, where* $N\Delta t \leq T$.

The proof of this theorem follows the ideas of a reformulation in the elliptic case (see [29] and [18]). The details are given in Appendix A.

4. A-priori error estimates.

In the following we are concerned with deriving an a-priori error estimate for our heterogeneous multiscale finite element method for advection-diffusion problems in the periodic setting. This estimate indicates the rates of convergence that we expect for the given HMM. The basic concept of this section will be similar to the one suggested in [31], chapter 12, for the treatment of the equation $\partial_t u - \triangle u = f$. Note that it is only the structure of the proof, which is still the same, but the completion is much more complicated. There are several additional difficulties in our problem, which are not treated in [31]. In particular the existence of time-dependent coefficient functions and the non-symmetric main part complicate the analysis. The problem of non-symmetry is treated in Lemma 4.7, where we give an equivalent formulation of the dual problem. Another novelty concerns the elliptic projection operator, which has to be introduced to finish the proof with an optimal order of convergence in space.

This rest of this section is structured as follows. First, we introduce a dual backward problem (Definition 4.2) which we use to derive an equation for the error (Lemma 4.4). After this, the contributions of the error identity need to be controlled by the L^2-norm of the error itself. These estimates are given in the Lemmas 4.6 and 4.8 below. For Lemma 4.8, it is essential to symmetrize the problem. This is achieved in Lemma 4.7

From now on, the error function between the homogenized solution and the HMM approximation is denoted by e^n, i.e.

$$e^n(\cdot) := u_H^n(\cdot) - u_0(t^n, \cdot)$$

for $0 \leq n \leq N$. With this notation, we now formulate the main result of this paper.

Theorem 4.1 (A-priori error estimate)**.** *Under the Assumptions 2.1 and if* v_H^0 *solves* $\int_{\mathbb{R}^d} v_H^0 \Phi_H^0 = \int_{\mathbb{R}^d} v_0 \Phi_H^0 \ \ \forall \Phi_H^0 \in V_H^1(\mathbb{R}^d)$*, the following a-priori error estimate is fulfilled:*

$$\begin{aligned} \|e^N\|_{L^2(\mathbb{R}^d)} \leq & CH^2 \max_{1\leq n\leq N} \left(\|u_0\|_{L^2(t^{n-1},t^n;H^2(\mathbb{R}^d))}\right) \left(\log \frac{t^{N+1}}{\triangle t}\right)^{\frac{1}{2}} \\ & + C\triangle t \max_{1\leq n\leq N} \left(\|\partial_t u_0\|_{L^2((t^{n-1},t^n)\times\mathbb{R}^d)}\right) \left(\log \frac{t^{N+1}}{\triangle t}\right)^{\frac{1}{2}} \\ & + C(\triangle t + H + h)\|(u_0, u_1)\|_{X^0(0,\bar{T})}. \end{aligned}$$

The theorem shows, that our multi-scale scheme is first order in time and second order in space. The term $(\triangle t + H + h)\|(u_0, u_1)\|_{X^0(0,\bar{T})}$ describes the approximation error determined by the qudrature rule. Choosing better approximations for the coefficient functions improves this error.

The rest of this section is concerned with proving Theorem 4.1. Before we start with introducing a suitable dual backward problem, we state a formulation of equation (4) which might be helpful, since the dual problem will be formulated in an analogous way.

Corollary 1. *Let $X_{\triangle t}$ be given by (13). For any solution of equation (4), we have*

$$G^N((u_0, u_1), (\Phi, \phi)) = (v_0, \Phi_+^0)_{L^2(\mathbb{R}^d)} \quad \forall (\Phi, \phi) \in X_{\triangle t}.$$

This result is obvious, since u_0 is continuous in t (which gives us $[u_0]_n = 0$) and since $(\partial_t u_0, \Phi)_{L^2(\mathbb{R}^d)} + E(t)((u_0, u_1), (\Phi, \phi)) = 0$. With regard to this corollary, we introduce the corresponding, discrete backward problem:

Definition 4.2 (Discrete backward problem)**.** We call $(z_H, z_h) \in \tilde{X}_H(0, t^{N+1})$ the solution of the *discrete dual backward problem*, if

$$G_H^N((\Phi_H, \phi_h), (z_H, z_h)) = ((\Phi_H)_-^N, e^N)_{L^2(\mathbb{R}^d)} \quad \forall\, (\Phi_H, \phi_h) \in X_H(0, t^N). \tag{15}$$

Remark 1. The discrete backward problem (15) is equivalent to the following backward Euler discretization. The initial value z_H^{N+1} is defined by $z_H^{N+1} := e^N$ and $(z_H^n, z_h^n) \in \tilde{I}_H$ is given by the equation

$$(\Phi_H, z_H^{n+1})_{L^2(\mathbb{R}^d)} = (\Phi_H, z_H^n)_{L^2(\mathbb{R}^d)} + \triangle t E_H^n((\Phi_H, \phi_h), (z_H^n, z_h^n)), \quad \forall\, (\Phi_H, \phi_h) \in I_H.$$

The following assumption is needed so that the error identity holds true.

Assumption 4.3. We assume that the discrete initial value v_H^0 is given by the following local L^2-projection $\int_{\mathbb{R}^d} v_H^0 \Phi_H^0 = \int_{\mathbb{R}^d} v_0 \Phi_H^0 \;\; \forall \Phi_H^0 \in V_H^1(\mathbb{R}^d)$.

Now, we are able to state an equation for the L^2-error $\|e^N\|_{L^2(\mathbb{R}^d)}$.

Lemma 4.4 (Error identity)**.** *Suppose that the assumptions 2.1 and 4.3 are fulfilled and that $(z_H, z_h) \in \tilde{X}_H(0, t^N)$ denotes the solution of the discrete backward problem (15). Then the following error identity holds true for all $(\Psi_H, \psi_h) \in X_H(0, t^N)$:*

$$\begin{aligned} \|e^N\|_{L^2(\mathbb{R}^d)}^2 &= \int_0^{t^N} E_H(t)((u_0 - \Psi_H, u_1 - \psi_h), (z_H, z_h))\, dt \\ &\quad + \sum_{n=1}^{N} (\Psi_H^n - u_0(t^n, \cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)} + \int_0^{t^N} (E - E_H)(t)((u_0, u_1), (z_H, z_h))\, dt. \end{aligned}$$

Remark 2. The error contributions on the right hand side of the error identity correspond to the space discretization, time discretization and data approximation errors. Estimates for these individual terms will be derived in the Lemmas 4.6 and 4.8 below.

Proof of Lemma 4.4. We start with the equation

$$\begin{aligned} (e^N, e^N)_{L^2(\mathbb{R}^d)} &= (u_H^N - u_0(t^N, \cdot), e^N)_{L^2(\mathbb{R}^d)} \\ &= (u_H^N - \Psi_H^N, e^N)_{L^2(\mathbb{R}^d)} + (\Psi_H^N - u_0(t^N, \cdot), e^N)_{L^2(\mathbb{R}^d)}. \end{aligned} \tag{16}$$

Using Assumption 4.3, Theorem 3.5 and Corollary 1, we have

$$G_H^N((u_H, \mathcal{K}_h(u_H)), (\Phi_H, \phi_h)) = G^N((u_0, u_1), (\Phi_H, \phi_h))$$

for all $(\Phi_H, \phi_h) \in X_H(0, t^N)$. Testing with (z_H, z_h) therefore yields

$$
\begin{aligned}
&(u_H^N - \Psi_H^N, e^N)_{L^2(\mathbb{R}^d)} = G_H^N((u_H - \Psi_H, \mathcal{K}_h(u_H) - \psi_h), (z_H, z_h)) \\
&= \quad G_H^N((u_H, \mathcal{K}_h(u_H)), (z_H, z_h)) - G_H^N((\Psi_H, \psi_h), (z_H, z_h)) \\
&= \quad G^N((u_0, u_1), (z_H, z_h)) - G_H^N((\Psi_H, \psi_h), (z_H, z_h)) \\
&= \quad (G^N - G_H^N)((u_0, u_1), (z_H, z_h)) + G_H^N((u_0 - \Psi_H, u_1 - \psi_h), (z_H, z_h)) \\
&= \quad \int_0^{t^N} (E - E_H)(t)((u_0, u_1), (z_H, z_h))\, dt + G_H^N((u_0 - \Psi_H, u_1 - \psi_h), (z_H, z_h)).
\end{aligned}
$$

This implies

$$
\begin{aligned}
(u_H^N - \Psi_H^N, e^N)_{L^2(\mathbb{R}^d)} &= \int_0^{t^N} (E - E_H)(t)((u_0, u_1), (z_H, z_h))\, dt \qquad (17) \\
&+ \sum_{n=0}^{N-1} \int_{t^n}^{t^{n+1}} (\partial_t(u_0 - \Psi_H^{n+1}), z_H^{n+1})_{L^2(\mathbb{R}^d)} + E_H(t)((u_0 - \Psi_H, u_1 - \psi_h), (z_H, z_h))\, dt \\
&+ \sum_{n=1}^{N-1} ([u_0 - \Psi_H]_n, z_H^{n+1})_{L^2(\mathbb{R}^d)} + (v_0 - \Psi_H^1, z_H^1)_{L^2(\mathbb{R}^d)}.
\end{aligned}
$$

Moreover, we see that

$$
\begin{aligned}
&\sum_{n=0}^{N-1} \left(\int_{t^n}^{t^{n+1}} (\partial_t(u_0 - \Psi_H^{n+1}), z_H^{n+1})_{L^2(\mathbb{R}^d)} + ([u_0 - \Psi_H]_n, z_H^{n+1})_{L^2(\mathbb{R}^d)} \right) + (v_0 - \Psi_H^1, z_H^1)_{L^2(\mathbb{R}^d)} \\
&= \quad \sum_{n=0}^{N-1} \left(\int_{t^n}^{t^{n+1}} (\partial_t u_0, z_H^{n+1})_{L^2(\mathbb{R}^d)} \right) - \sum_{n=1}^{N-1} ([\Psi_H]_n, z_H^{n+1})_{L^2(\mathbb{R}^d)} + (v_0 - \Psi_H^1, z_H^1)_{L^2(\mathbb{R}^d)} \\
&= \quad \sum_{n=0}^{N-1} (u_0(t^{n+1}, \cdot) - u_0(t^n, \cdot), z_H^{n+1})_{L^2(\mathbb{R}^d)} - \sum_{n=1}^{N-1} (\Psi_H^{n+1} - \Psi_H^n, z_H^{n+1})_{L^2(\mathbb{R}^d)} \\
&\quad + (v_0 - \Psi_H^1, z_H^1)_{L^2(\mathbb{R}^d)} \\
&= \quad \sum_{n=0}^{N-1} (u_0(t^{n+1}, \cdot), z_H^{n+1})_{L^2(\mathbb{R}^d)} - \sum_{n=1}^{N-1} (u_0(t^n, \cdot), z_H^{n+1})_{L^2(\mathbb{R}^d)} \\
&\quad - \sum_{n=0}^{N-1} (\Psi_H^{n+1}, z_H^{n+1})_{L^2(\mathbb{R}^d)} + \sum_{n=1}^{N-1} (\Psi_H^n, z_H^{n+1})_{L^2(\mathbb{R}^d)} \\
&= \quad \sum_{n=1}^{N-1} (\Psi_H^n - u_0(t^n, \cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)} + (u_0(t^N, \cdot) - \Psi_H^N, z_H^N)_{L^2(\mathbb{R}^d)}
\end{aligned}
$$

Combining this result with (16), we obtain:

$$
\begin{aligned}
&\|e^N\|_{L^2(\mathbb{R}^d)}^2 = (u_H^N - \Psi_H^N, e^N)_{L^2(\mathbb{R}^d)} + (\Psi_H^N - u_0(t^N, \cdot), e^N)_{L^2(\mathbb{R}^d)} \\
&= \quad \int_0^{t^N} (E - E_H)(t)((u_0, u_1), (z_H, z_h))\, dt + \sum_{n=1}^{N-1} (\Psi_H^n - u_0(t^n, \cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)} \\
&\quad + (u_0(t^N, \cdot) - \Psi_H^N, z_H^N)_{L^2(\mathbb{R}^d)} + (\Psi_H^N - u_0(t^N, \cdot), e^N)_{L^2(\mathbb{R}^d)} \\
&\quad + \sum_{n=0}^{N-1} \int_{t^n}^{t^{n+1}} E_H^{n+1}((u_0 - \Psi_H, u_1 - \psi_h), (z_H, z_h))\, dt
\end{aligned}
$$

$$= \int_0^{t^N} (E - E_H)(t)((u_0, u_1), (z_H, z_h))\, dt + \sum_{n=1}^{N} (\Psi_H^n - u_0(t^n, \cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)}$$
$$+ \int_0^{t^N} E_H(t)((u_0 - \Psi_H, u_1 - \psi_h), (z_H, z_h))\, dt.$$

□

In the following we derive some estimates, which we need to control the right hand side of the error identity in Lemma 4.4. For simplification, we introduce the following notations.

Definition 4.5. We define:

$$\bar{u}_0^{n+1} := \frac{1}{\triangle t} \int_{t^n}^{t^{n+1}} u_0(t, \cdot)\, dt,$$
$$\tilde{\mathcal{I}}_H(u_0)(\cdot, x)_{|(t^n, t^{n+1}] \times \mathbb{R}^d} := \mathcal{I}_H(\bar{u}_0^{n+1})(x)$$

Here $\mathcal{I}_H$ denotes the corresponding Lagrange interpolation operator. Moreover, we establish the notation $\tilde{z}_H^n := (z_H^n, z_h^n)$.

In the next lemma, we derive an estimate for the space derivative contributions of the solution of the discrete backward problem in Definition 4.2:

Lemma 4.6. *Assume that the general assumptions 2.1 are fulfilled, then we have the following estimate for the solution of the discrete backward problem (15):*

$$\sum_{n=0}^{N} \triangle t \left(\|z_H^n\|^2_{H^1(\mathbb{R}^d)} + \|z_h^n\|^2_{L^2(\mathbb{R}^d, H^1(Y))} \right) \le C \|e^N\|^2_{L^2(\mathbb{R}^d)}.$$

Proof.

Since $(z_H^n, z_H^{n+1})_{L^2(\mathbb{R}^d)} = (z_H^n, z_H^n)_{L^2(\mathbb{R}^d)} + \triangle t E_H^n(\tilde{z}_H^n, \tilde{z}_H^n) \ge \|z_H^n\|^2_{L^2(\mathbb{R}^d)}$, we have

$$\|z_H^0\|_{L^2(\mathbb{R}^d)} \le \|z_H^1\|_{L^2(\mathbb{R}^d)} \le ... \le \|z_H^{N+1}\|_{L^2(\mathbb{R}^d)}. \tag{18}$$

This implies $\sum_{n=0}^{N} \triangle t \|z_H^n\|^2_{L^2(\mathbb{R}^d)} \le t^{N+1} \|e^N\|^2_{L^2(\mathbb{R}^d)}$. (19)

Using the Poincare inequality for functions with mean zero, we get

$$\sum_{n=0}^{N} \triangle t \|z_h^n\|^2_{L^2(\mathbb{R}^d, H^1(Y))} \le C \sum_{n=0}^{N} \triangle t |z_h^n|^2_{L^2(\mathbb{R}^d, H^1(Y))}. \tag{20}$$

(19) and (20) imply that it is sufficient to bound the semi-norms in the estimate of Lemma 4.6, i.e. we restrict ourself to the following term:

$$\sum_{n=0}^{N} \triangle t \left(|z_H^n|^2_{H^1(\mathbb{R}^d)} + |z_h^n|^2_{L^2(\mathbb{R}^d, H^1(Y))} \right).$$

First we get

$$
\begin{aligned}
(z_H^N, e^N)_{L^2(\mathbb{R}^d)} &= \sum_{n=0}^{N} \triangle t E_H^n(\tilde{z}_H^n, \tilde{z}_H^n) + \sum_{n=0}^{N-1} (z_H^{n+1} - z_H^n, z_H^{n+1})_{L^2(\mathbb{R}^d)} + \|z_H^0\|^2_{L^2(\mathbb{R}^d)} \\
&= \sum_{n=0}^{N} \triangle t E_H^n(\tilde{z}_H^n, \tilde{z}_H^n) + \frac{1}{2} \sum_{n=0}^{N-1} (z_H^{n+1} - z_H^n, z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)} \\
&\quad + \frac{1}{2} \sum_{n=0}^{N-1} (z_H^{n+1}, z_H^{n+1})_{L^2(\mathbb{R}^d)} - \frac{1}{2} \sum_{n=0}^{N-1} (z_H^n, z_H^n)_{L^2(\mathbb{R}^d)} + \|z_H^0\|^2_{L^2(\mathbb{R}^d)} \\
&\geq \sum_{n=0}^{N} \triangle t E_H^n(\tilde{z}_H^n, \tilde{z}_H^n).
\end{aligned}
$$

This inequality together with the ellipticity of E_H^n (with constant 1) yields

$$
\begin{aligned}
\sum_{n=0}^{N} \triangle t \left(|z_H^n|^2_{H^1(\mathbb{R}^d)} + |z_h^n|^2_{L^2(\mathbb{R}^d, H^1(Y))} \right) &\leq \sum_{n=0}^{N} \triangle t E_H^n(\tilde{z}_H^n, \tilde{z}_H^n) \\
&\leq (z_H^N, e^N)_{L^2(\mathbb{R}^d)} \leq \|e^N\|^2_{L^2(\mathbb{R}^d)}.
\end{aligned}
$$

□

The bilinear form E_H^n is not symmetric, which complicates the analysis. To avoid rather technical estimates to treat the non-symmetric case, we use a symmetrization result which is given in Lemma 4.7 below. It is well known, that the standard homogenization of problems like (3), yields a limit problem of the type

$$
\partial_t u_0 - \nabla \cdot (\bar{A} \nabla u_0) = 0,
$$

where $\bar{A}$ is a symmetric, coercive diffusion matrix, only depending on t (see for instance Allaire and Raphael [7]). The relation between this problem and the two-scale homogenized equation was shown in [17], Theorem 3.8. The following lemma is the equivalent result in the discrete setting. It simplifies the subsequent analysis enormously.

Lemma 4.7. *We introduce the operator $T_h^n : V_H^1(\mathbb{R}^d) \to V_H^0(\mathbb{R}^d, W_h^1(Y))$, where $T_h^n(\Phi_H) \in V_H^0(\mathbb{R}^d, W_h^1(Y))$ is the unique solution of*

$$
\begin{aligned}
&\int_Y \left(b_h(t^n, y) - k_h(t^n, y) \bar{b}_h(t^n) \right) \cdot \nabla_x \Phi_H(x) \;\; \phi_h(y) \;\; dy \\
&\quad = \int_Y A_h(t^n, y) \nabla_y \phi_h(y) \cdot (\nabla_x \Phi_H(x) + \nabla_y T_h^n(\Phi_H)(x, y)) \;\; dy \\
&\quad\quad + \int_Y b_h(t^n, y) \cdot \nabla_y \phi_h(y) \;\; T_h^n(\Phi_H)(x, y) \;\; dy \qquad \text{for all } \phi_h \in W_h^1(Y).
\end{aligned}
$$

Moreover, we define the symmetric bilinearform $S_H^n : V_H^1(\mathbb{R}^d) \times V_H^1(\mathbb{R}^d) \to \mathbb{R}$ by:

$$
\begin{aligned}
&S_H^n(\Phi_H, \Psi_H) \\
&\quad := \int_{\mathbb{R}^d} \int_Y A_h(t^n, y)(\nabla_x \Phi_H(x) + \nabla_y T_h^n(\Phi_H)(x, y))(\nabla_x \Psi_H(x) + \nabla_y T_h^n(\Psi_H)(x, y)) \, dy \, dx.
\end{aligned}
$$

With theses definitions, the solution of the discrete backward problem (15) (see also Remark 1) fulfills the equation

$$
(\Phi_H, z_H^{n+1})_{L^2(\mathbb{R}^d)} = (\Phi_H, z_H^n)_{L^2(\mathbb{R}^d)} + \triangle t S_H^n(\Phi_H, z_H^n), \tag{21}
$$

for all $\Phi_H \in V_H^1(\mathbb{R}^d)$*. We also have* $T_h^n(z_H^n) = z_h^n$ *and*

$$|\Phi_H|^2_{H^1(\mathbb{R}^d)} \leq S_H^n(\Phi_H, \Phi_H) \text{ for all } \Phi_H \in V_H^1.$$

Proof. The proof of this lemma is completely analogous to the proof in the continuous setting, which can be found in [17], Theorem 3.8. □

Lemma (4.8) below is an estimate for the contributions of the discrete time derivatives of the solution of the discrete backward problem in Definition 4.2. To prove it, we make use of the symmetrisation result in Lemma 4.7. On the basis of this result, the proof is quite analogous to the one presented in the book of Thomée [31], chapter 12.

Lemma 4.8. *Under the general assumptions 2.1, we have the following estimate for the solution* $(z_H, z_h) \in \tilde{X}_H(0, t^{N+1})$ *of the discrete backward problem (15):*

$$\sum_{n=0}^{N} \frac{1}{\sqrt{\triangle t}} \|z_H^{n+1} - z_H^n\|_{L^2(\mathbb{R}^d)} \leq C \left(\log \frac{t^{N+1}}{\triangle t} \right)^{\frac{1}{2}} \|e^N\|_{L^2(\mathbb{R}^d)}.$$

The proof of Lemma 4.8 is given in Appendix B.

With the preceding estimates, we are now ready to prove the a-priori error estimate of Theorem 4.1.

Proof of Theorem 4.1. Choosing $(\Psi_H, \psi_h) = (\tilde{\mathcal{I}}_H(u_0), 0)$ in Lemma 4.4 yields:

$$\begin{aligned} \|e^N\|^2_{L^2(\mathbb{R}^d)} &= \int_0^{t^N} E_H(t)((u_0 - \tilde{\mathcal{I}}_H(u_0), u_1), (z_H, z_h))\, dt \qquad (22) \\ &+ \int_0^{t^N} (E - E_H)(t)((u_0, u_1), (z_H, z_h))\, dt + \sum_{n=1}^{N} (\tilde{\mathcal{I}}_H(u_0)^n - u_0(t^n, \cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)}. \end{aligned}$$

Now we estimate the different summands. Since we have Lipschitz continuous coefficient functions we use Lemma 4.6 to get

$$\begin{aligned} &\int_0^{t^N} (E - E_H)(t)((u_0, u_1), (z_H, z_h))\, dt \\ &\leq C(\triangle t + H + h) \|(u_0, u_1)\|_{X^0(0,\bar{T})} \left(\sum_{n=1}^{N} \triangle t \left(\|z_H^n\|^2_{H^1(\mathbb{R}^d)} + \|z_h^n\|^2_{L^2(\mathbb{R}^d, H^1(Y))} \right) \right)^{\frac{1}{2}} \\ &\leq C(\triangle t + H + h) \|(u_0, u_1)\|_{X^0(0,\bar{T})} \|e^N\|_{L^2(\mathbb{R}^d)}. \end{aligned}$$

The third summand in the right hand side of (22) is separated as follows:

$$\begin{aligned} &\sum_{n=1}^{N} (\tilde{\mathcal{I}}_H(u_0)^n - u_0(t^n, \cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)} \\ &= \sum_{n=1}^{N} (\tilde{\mathcal{I}}_H(u_0)^n - \bar{u}_0^n, z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)}) + \sum_{n=1}^{N} (\bar{u}_0^n - u_0(t^n, \cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)}. \end{aligned}$$

Since

$$\begin{aligned} \|\bar{u}_0^n - u_0(t^n, \cdot)\|^2_{L^2(\mathbb{R}^d)} &\leq \int_{\mathbb{R}^d} \left(\frac{1}{\triangle t} \int_{t^{n-1}}^{t^n} \int_t^{t^n} |\partial_t u_0(s, \cdot)|\, ds\, dt \right)^2 \\ &\leq \triangle t \|\partial_t u_0\|^2_{L^2((t^{n-1}, t^n) \times \mathbb{R}^d)} \end{aligned}$$

we get by means of Lemma 4.8

$$\begin{aligned}
&\sum_{n=1}^{N}(\bar{u}_0^n - u_0(t^n,\cdot), z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)} \\
&\leq \triangle t \max_{1\leq n\leq N}\left(\|\partial_t u_0\|_{L^2((t^{n-1},t^n)\times\mathbb{R}^d)}\right)\sum_{n=0}^{N}\frac{1}{\sqrt{\triangle t}}\|z_H^{n+1} - z_H^n\|_{L^2(\mathbb{R}^d)} \\
&\leq C\triangle t \max_{1\leq n\leq N}\left(\|\partial_t u_0\|_{L^2((t^{n-1},t^n)\times\mathbb{R}^d)}\right)\left(\log\frac{t^{N+1}}{\triangle t}\right)^{\frac{1}{2}}\|e^N\|_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

And moreover:

$$\begin{aligned}
&\sum_{n=1}^{N}(\tilde{\mathcal{I}}_H(u_0)^n - \bar{u}_0^n, z_H^{n+1} - z_H^n)_{L^2(\mathbb{R}^d)}) \\
&\leq C\sum_{n=1}^{N}H^2\|\bar{u}_0^n\|_{H^2(\mathbb{R}^d)}\|z_H^{n+1} - z_H^n\|_{L^2(\mathbb{R}^d)} \\
&\leq C\sum_{n=1}^{N}H^2\frac{1}{\sqrt{\triangle t}}\|u_0\|_{L^2(t^{n-1},t^n;H^2(\mathbb{R}^d))}\|z_H^{n+1} - z_H^n\|_{L^2(\mathbb{R}^d)} \\
&\leq CH^2 \max_{1\leq n\leq N}\left(\|u_0\|_{L^2(t^{n-1},t^n;H^2(\mathbb{R}^d))}\right)\left(\log\frac{t^{N+1}}{\triangle t}\right)^{\frac{1}{2}}\|e^N\|_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

It remains to estimate the first summand in the right hand side of (22). To do so, we denote $\beta := \sup_{0\leq t\leq T}\|E(t)\|_{\mathcal{L}(I,I')}$. Moreover, we define the projection operator $\mathcal{P}_{H,h}^n = (P_H^n, P_h^n) : I_0 \to \tilde{I}_H$ by:

$$E_H^n(v,\Phi_{H,h}) = E_H^n(\mathcal{P}_{H,h}^n(v),\Phi_{H,h}) \;\; \forall\Phi_{H,h}\in\tilde{I}_H \;\text{ and }\; \forall v\in I_0. \tag{23}$$

$\mathcal{P}_{H,h}^n$ is well-defined due to the Lax-Milgram Theorem. Since for $v = (v_0, v_1) \in I_0$

$$\|\mathcal{P}_{H,h}^n(v)\|_{I_0}^2 \leq E_H^n(\mathcal{P}_{H,h}^n(v),\mathcal{P}_{H,h}^n(v)) = E_H^n(v,\mathcal{P}_{H,h}^n(v)) \leq \beta\|\mathcal{P}_{H,h}^n(v)\|_{I_0}\|v\|_{I_0}$$

and since

$$\begin{aligned}
\|P_H^n\|_{\mathcal{L}(H^1(\mathbb{R}^d),V_H^1)} &= \sup_{v_0\in H^1(\mathbb{R}^d)\setminus\{0\}}\frac{|P_H^n v_0|_{H^1(\mathbb{R}^d)}}{|v_0|_{H^1(\mathbb{R}^d)}} \\
&\leq \sup_{(v_0,v_1)\in I_0\setminus\{0\}}\frac{|P_H^n v_0|_{H^1(\mathbb{R}^d)} + |P_h^n v_1|_{L^2(\mathbb{R}^d,H^1(Y))}}{|v_0|_{H^1(\mathbb{R}^d)} + |v_1|_{L^2(\mathbb{R}^d,H^1(Y))}},
\end{aligned}$$

we also get:

$$\|P_H^n\|_{\mathcal{L}(H^1(\mathbb{R}^d),V_H^1)} \leq \|\mathcal{P}_{H,h}^n\|_{\mathcal{L}(I_0,\tilde{I}_H)} \leq \beta. \tag{24}$$

Now, we are prepared to estimate the first summand in the right hand side of (22). Using Remark 1, (23) and (24) we get:

$$\begin{aligned}
&\int_0^{t^N} E_H(t)((u_0 - \tilde{\mathcal{I}}_H(u_0), u_1), (z_H, z_h))\,dt \\
&= \sum_{n=1}^{N} E_H^n((\int_{t^{n-1}}^{t^n} u_0 - \tilde{\mathcal{I}}_H(u_0)\,dt, \int_{t^{n-1}}^{t^n} u_1\,dt), (z_H^n, z_h^n)) \\
&= \sum_{n=1}^{N} E_H^n((P_H^n\left(\int_{t^{n-1}}^{t^n} u_0 - \tilde{\mathcal{I}}_H(u_0)\,dt\right), P_h^n\left(\int_{t^{n-1}}^{t^n} u_1\,dt\right)), (z_H^n, z_h^n)) \\
&= \sum_{n=1}^{N} (\frac{z_H^{n+1} - z_H^n}{\triangle t}, P_H^n\left(\int_{t^{n-1}}^{t^n} u_0 - \tilde{\mathcal{I}}_H(u_0)\,dt\right))_{L^2(\mathbb{R}^d)} \\
&\leq \sum_{n=1}^{N} \|\frac{z_H^{n+1} - z_H^n}{\triangle t}\|_{L^2(\mathbb{R}^d)} \|P_H^n\|_{\mathcal{L}(H^1(\mathbb{R}^d), V_H^1)} H^2 \|\int_{t^{n-1}}^{t^n} u_0\,dt\|_{H^2(\mathbb{R}^d)} \\
&\leq \sum_{n=1}^{N} \|\frac{z_H^{n+1} - z_H^n}{\triangle t}\|_{L^2(\mathbb{R}^d)} \beta H^2 \sqrt{\triangle t} \|u_0\|_{L^2((t^{n-1}, t^n), H^2(\mathbb{R}^d))} \\
&\leq \beta H^2 \max_{1 \leq n \leq N} \left(\|u_0\|_{L^2(t^{n-1}, t^n; H^2(\mathbb{R}^d))}\right) \left(\log \frac{t^{N+1}}{\triangle t}\right)^{\frac{1}{2}} \|e^N\|_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

□

5. **Numerical experiments.** In the following we look at two model problems to demonstrate the applicability of the HMM given by Definition 3.3. In the first example, we apply the method to an advection-diffusion problem with a non-zero drift. The time-dependent coefficients are periodic in space. Here, the exact solution u^ϵ is unknown, but since $u_0(t, x - \frac{B(t)}{\epsilon})$ is a good approximation of u^ϵ, we use this as a reliable reference. u_0 can be determined very efficiently by using the associated homogenized macro problem, see Theorem 4.7 in [17]. Here, we need to solve the two corresponding cell problems for every time step and the resulting macro problem afterwards. In the second example we will apply the method to an advection-diffusion problem without drift, but with a heterogeneous diffusion matrix. Here the standard homogenization theory fails, so that we have to determine u^ϵ by a standard computation on a very fine grid. We will see that to obtain a sufficiently accurate approximation with a Backward-Euler (Linear-)Finite-Element Scheme (BE-FES), the grid needs to be about 6 times finer than for a comparable approximation computed with the HMM.

In this chapter we will use the following notations: For the n'th time step u_H^n denotes the HMM approximation, whereas u_{BWS}^n denotes the approximation gained by a Backward-Euler Finite-Element Scheme. The corresponding error functions are given by $e^n := u_0(t^n, \cdot) - u_H^n$ and $e_{BWS}^n := u_0(t^n, \cdot - \frac{B(t^n)}{\epsilon}) - u_{BWS}^n$. $\|e^n\|_{L^2(\Omega)}^{rel} := \frac{\|e^n\|_{L^2(\Omega)}}{\|u_0(t^n, \cdot)\|_{L^2(\Omega)}}$ and $\|e_{BWE}^n\|_{L^2(\Omega)}^{rel} := \frac{\|e_{BWE}^n\|_{L^2(\Omega)}}{\|u_0(t^n, \cdot - \frac{B(t^n)}{\epsilon})\|_{L^2(\Omega)}}$ denote the associated relative errors. N will define the maximal number of time steps, i.e. $t^N = T$ or alternatively $\triangle t = \frac{\bar{T}}{N}$, with time step size $\triangle t$. In both numerical experiments the observed time interval $[0, \bar{T}]$ will be given by $[0, 0.3]$ and the fineness parameter ϵ will be set to $\epsilon = 0.01$. For $k, m \in \mathbb{N}_+$, $\delta \in \mathbb{R}_+^m$ and the error function $g : \mathbb{R}_+^m \to \mathbb{R}_+$

TABLE 1. *Model problem 1. Relative errors for $N = 20$ and decreasing mesh sizes (H, h).*

H	h	$\|e^1\|^{rel}_{L^2(\Omega)}$	$\|e^N\|^{rel}_{L^2(\Omega)}$	$\|e^1_{BWE}\|^{rel}_{L^2(\Omega)}$	$\|e^N_{BWE}\|^{rel}_{L^2(\Omega)}$
$2^{-\frac{5}{2}}$	$2^{-\frac{5}{2}}$	0.497	2.674	0.349	1.76
2^{-3}	2^{-3}	0.339	0.727	0.292	1.17
$2^{-\frac{7}{2}}$	$2^{-\frac{7}{2}}$	0.206	0.297	0.246	0.983
2^{-4}	2^{-4}	0.1748836	0.1945293	0.292	0.955
2^{-4}	2^{-5}	0.1748802	0.1945079	0.292	0.955
2^{-4}	2^{-6}	0.1748784	0.1945025	0.292	0.955
$2^{-\frac{9}{2}}$	$2^{-\frac{7}{2}}$	0.1428528	0.1372253	0.318	0.968
$2^{-\frac{9}{2}}$	$2^{-\frac{9}{2}}$	0.1428393	0.1371633	0.318	0.968
$2^{-\frac{9}{2}}$	$2^{-\frac{11}{2}}$	0.1428358	0.1371477	0.318	0.968
2^{-5}	2^{-5}	0.135	0.112	0.334	0.967

we define the experimental order of convergence (EOC) of g in $(k\delta \to \delta)$ by

$$\mathrm{EOC}_{(k\delta\to\delta)}(g) := \frac{\log\left(\frac{g(k\delta)}{g(\delta)}\right)}{\log(k)}. \tag{25}$$

Model problem 1. In our first numerical test, we look at the following model problem: find $u^\epsilon \in L^2(0, 0.3; H^1(\mathbb{R}^2))$, with

$$\partial_t u^\epsilon - \nabla \cdot \left(A\Big(t, \frac{x}{\epsilon}\Big)\nabla u^\epsilon\right) + \epsilon^{-1} b\Big(t, \frac{x}{\epsilon}\Big) \cdot \nabla u^\epsilon = 0 \ \text{ in } \ (0, 0.3) \times \mathbb{R}^2 \ \text{ and}$$

$$u^\epsilon(0, x_1, x_2) = \begin{cases} 5\sin(5\pi x_1)\sin(5\pi x_2) & \text{in } [-0.2, 0.2]^2 \\ 0, & \text{else.} \end{cases}$$

Here $A^\epsilon(t, x) = A(t, \frac{x}{\epsilon})$ and $b^\epsilon(t, x) = b(\frac{x}{\epsilon})$ are defined by:

$$A(t, y) := \begin{pmatrix} 1 + \frac{t}{2}\sin(2\pi y_1)\cos(2\pi y_2) & 0 \\ 0 & 1 + \frac{t}{2}\sin(2\pi y_1)\cos(2\pi y_2) \end{pmatrix} \ \text{ and}$$

$$b(t, y) := \begin{pmatrix} -\sin(2\pi y_1)\cos(2\pi y_2) + \frac{1}{10} \\ \cos(2\pi y_1)\sin(2\pi y_2) - \frac{1}{10} \end{pmatrix}.$$

In Table 1 we see, that the relative error between the homogenized solution and the approximation gained by the HMM in Definition 3.3 is small and diminishing for decreasing mesh size, whereas the relative error between $u_0(t^N, \cdot - \frac{B(t^N)}{\epsilon})$ and the Backward-Euler Finite-Element solution u^N_{BWS} remains essentially the same for the whole computation series. We do not observe any convergence for the BE-FE Scheme for these refinement levels. In fact, a mesh size of at least $H^{-\frac{13}{2}}$ (roughly ϵ) is required, so that the BE-FE Scheme yields a first reliable approximation of u^ϵ in Model Problem 1. Table 1 also gives a hint for the relation between macro mesh size H and micro mesh size h. In this example we observe that choosing h smaller than H, has almost no effect on the quality of the solution. The computation time for solving one cell problem is increased, but the error remains basically the same. For small values of H it may be even expedient to choose h larger than H, since, in this case, there is no need for solving the cell problems with the same accuracy. Compare for instance the computations for $(H, h) = (2^{-\frac{9}{2}}, 2^{-\frac{7}{2}})$ and $(H, h) = (2^{-\frac{9}{2}}, 2^{-\frac{11}{2}})$. In Figure 1 we give a comparison between the isolines of

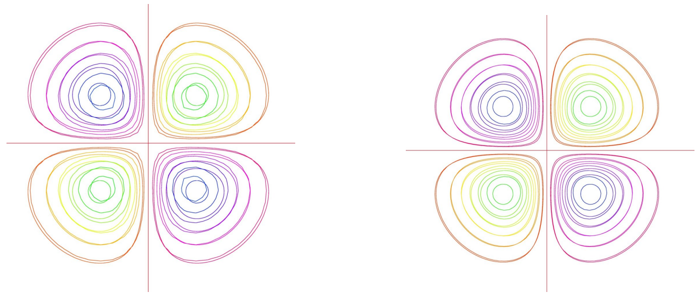

FIGURE 1. *Model problem 1. Comparison of the isolines between* $u_0(t^N,\cdot)$ *and* u_H^N *for different resolution of the HMM* ($(H,h) = (2^{-4},2^{-5})$, $N = 20$ *(left), and* $(H,h) = (2^{-5},2^{-6})$, $N = 40$ *(right)). The color gradient is from green (minimum* $-4.51\cdot 10^{-3}$*) to blue (maximum* $4.51\cdot 10^{-3}$*).*

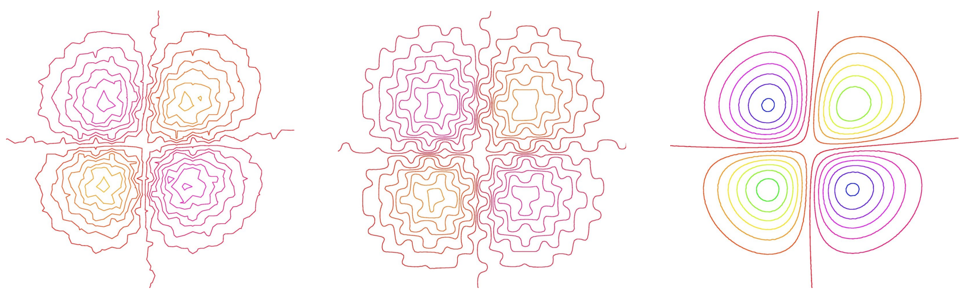

FIGURE 2. *Model problem 1. The isolines of approximations of* $u^\epsilon(t^N,\cdot)$ *are shown for a Backward Euler FE Scheme with different space and time resolution (*$H = 2^{-4}$, $N = 20$ *(left),* $H = 2^{-6}$, $N = 40$ *(middle), and* $H = 2^{-\frac{13}{2}}$, $N = 40$ *(left)). The color gradient is from green (minimum* $-4.51\cdot 10^{-3}$*) to blue (maximum* $4.51\cdot 10^{-3}$*).*

TABLE 2. *Model problem 1. With Definition* (25)*, we calculate the following EOC's in space, i.e.* $EOC_{(2(H,h)\to(H,h))}$*. For each computation* N *is fixed (and therefore also* $\triangle t$*). To get reliable results for the convergence rate in space, we choose* N *large enough.*

N	$k*H\to H$	$k*h\to h$	EOC(e^N)	EOC(e^N_{BWE})
10	$2^{-\frac{5}{2}}\to 2^{-\frac{7}{2}}$	$2^{-\frac{5}{2}}\to 2^{-\frac{7}{2}}$	1.9105	0.8155
20	$2^{-3}\to 2^{-4}$	$2^{-3}\to 2^{-4}$	1.9015	0.2929
40	$2^{-\frac{7}{2}}\to 2^{-\frac{9}{2}}$	$2^{-\frac{7}{2}}\to 2^{-\frac{9}{2}}$	2.0247	0.0204
40	$2^{-4}\to 2^{-5}$	$2^{-4}\to 2^{-5}$	1.618	-0.0203

the exact solution and the HMM solution, gained for different values of $(\triangle t,H,h)$. We see that the lines match quite well. The better correspondence of the isolines is achieved for higher resolution level of the computational grid. In Figure 2, on the other hand, the isolines of BE-FES approximations at different refinement levels are expressed. Obviously, the first two approximations (for $(\triangle t,H) = (\frac{3}{2}\cdot 10^{-3},2^{-4})$

TABLE 3. *Model problem 1. With Definition (25), we calculate the following EOC's in time, i.e. $EOC_{(2\triangle t \to \triangle t)}$. For each computation (H, h) is fixed. We decrease (H, h) with $\triangle t$ to reduce the influence of the error in space. For $10 \to 20$ and $H = 2^{-\frac{9}{2}}$, the mesh size does not seem to be small enough. We observe the EOC's for $N \to 2N$ with $N = 3, 5, 10$.*

$N \to k \cdot N$	H	h	EOC(e^N)	EOC(e^N_{BWE})
$3 \to 6$	2^{-4}	2^{-5}	1.1164	-0.2045
$5 \to 10$	$2^{-\frac{9}{2}}$	$2^{-\frac{11}{2}}$	1.0425	-0.1179
$10 \to 20$	$2^{-\frac{9}{2}}$	$2^{-\frac{11}{2}}$	0.6608	-0.0497
$10 \to 20$	2^{-5}	2^{-6}	0.8994	-0.0488

and $(\triangle t, H) = (\frac{3}{4} \cdot 10^{-3}, 2^{-6}))$ are not reliable. Comparing them with the isolines of the exact solution in Figure 1, we immediately verify that they significantly differ in shape and height. Only for $(\triangle t, H) = (\frac{3}{4} \cdot 10^{-3}, 2^{-\frac{13}{2}})$ the solution is reasonable.

Experimental orders of convergence are shown in Table 2 and Table 3. In order to get reliable results for the EOC for the space refinement, we needed to choose $\triangle t$ small in comparison to H^2 (see Theorem 4.1). Therefore, taking only values fulfilling $\triangle t \ll H^2$, we tried to assure that the influence of $\triangle t$ is kept small. With regard to Theorem 4.1, we expect an EOC of 2, if u^ϵ is a regular solution and if we have good approximations of the coefficient functions. This is confirmed by the table. The relatively bad value (1.618) for $2^{-4} \to 2^{-5}$ is probably due to $\triangle t$ no yet being small enough. Moreover, we point out that the EOC's for the BW-FE Scheme directly imply that we do not have a convergence to u^ϵ, as long as we do not have a highly refined grid that captures ϵ, i.e.$H \approx \epsilon = 0.01$.

In Table 3, corresponding time EOC's are shown. For the BW-FE Scheme it is obvious that we cannot observe a convergence in time as long as we are not fine enough in space. For the HMM we notice, that the time convergence seems to be linear, which also corresponds with Theorem 4.1. Again we needed to guarantee that we have a sufficiently small mesh size in order to avoid that it has a visible influence on the results. For this purpose we assumed that roughly $H^2 \leq \triangle t$ holds true. For $(\triangle t, H) = (\frac{\bar{T}}{20}, 2^{-\frac{9}{2}})$, this assumption is not fulfilled. Immediately we see the loss of quality at this result.

Model problem 2. In the second numerical test, we observe the following model problem: find $u^\epsilon \in L^2(0, 0.3; \mathring{H}^1([-1, 1]^2))$ with

$$\begin{aligned} \partial_t u^\epsilon - \nabla \cdot (A^\epsilon(x) \nabla u^\epsilon) + \epsilon^{-1} b^\epsilon(t, x) \cdot \nabla u^\epsilon = 0 \ \text{ in } \ (-1, 1)^2 \times (0, 0.3), \\ u^\epsilon = 0 \ \text{ on } \ \partial(-1, 1)^2 \times (0, \bar{T}) \ \text{ and} \\ u^\epsilon(0, x_1, x_2) = \begin{cases} 10 \sin(2\pi x_1) \sin(2\pi x_2) & \text{in } [0, 0.5]^2 \\ 0, & \text{else.} \end{cases} \end{aligned}$$

Here $b^\epsilon(t, x) = b(t, \frac{x}{\epsilon})$ is defined by

$$b(t, y) := \begin{pmatrix} \frac{3}{2} \sin(2\pi y_1) \sin(2\pi y_2) \\ \frac{3}{2} \cos(2\pi y_1) \cos(2\pi y_2) \end{pmatrix}$$

TABLE 4. *Model problem 2. Relative errors for* $N = 20$ *and decreasing mesh sizes. gain we see that the relative HMM error is significantly decreasing, whereas the relative BE-FES error for* $T = 0.3$ *remains at a value of roughly* 0.5.

H	h	$\|e^1\|^{rel}_{L^2(\Omega)}$	$\|e^N\|^{rel}_{L^2(\Omega)}$	$\|e^1_{BWE}\|^{rel}_{L^2(\Omega)}$	$\|e^N_{BWE}\|^{rel}_{L^2(\Omega)}$
$2^{-\frac{5}{2}}$	$2^{-\frac{5}{2}}$	0.153	0.49	0.245	1.198
2^{-3}	2^{-3}	0.086	0.209	0.137	0.36
$2^{-\frac{7}{2}}$	$2^{-\frac{7}{2}}$	0.047	0.101	0.172	0.551
2^{-4}	2^{-4}	0.031	0.049	0.192	0.533
$2^{-\frac{9}{2}}$	$2^{-\frac{9}{2}}$	0.024	0.035	0.183	0.447
2^{-5}	2^{-5}	0.02	0.027	0.192	0.459

and A^ϵ by

$$A^\epsilon(x) := \begin{pmatrix} a(\frac{x}{\epsilon}) + \frac{1}{5}\sin(2\pi\frac{x_1}{\epsilon})^2 & 0 \\ 0 & a(\frac{x}{\epsilon}) + \frac{1}{5}\sin(2\pi\frac{x_1}{\epsilon})^2 \end{pmatrix}$$

where

$$a(y) := \begin{cases} 1+\frac{1}{2}\sin(2\pi y_1)\sin(2\pi y_2) & \text{on } [-0.1,0.1]^2, \\ 1+\frac{1}{5}\sin(4\pi y_1)(\cos(2\pi y_2)-1) & \text{on } [-0.2,0.2]^2\setminus[-0.1,0.1]^2, \\ 1+\frac{1}{10}\left(\cos(2\pi y_1)-1\right)\left(\cos(4\pi y_2)-1\right) & \text{on } [-0.3,0.3]^2\setminus[-0.2,0.2]^2, \\ 1+(\frac{3}{10}\sin(2\pi y_1)\sin(2\pi y_1)(\cos(8\pi y_2)-1)) & \text{on } [-0.4,0.4]^2\setminus[-0.3,0.3]^2, \\ 1+(\frac{1}{2}\sin(4\pi y_1)\sin(2\pi y_2)) & \text{on } [-0.5,0.5]^2\setminus[-0.4,0.4]^2, \\ 1+(\frac{1}{5}\sin(\pi y_1)(\cos(2\pi y_2)-1)) & \text{on } [-0.6,0.6]^2\setminus[-0.5,0.5]^2, \\ 1+(\frac{1}{10}(\cos(4\pi y_1)-1)(\cos(2\pi y_2)-1)) & \text{on } [-0.7,0.7]^2\setminus[-0.6,0.6]^2, \\ 1+(\frac{3}{10}\sin(2\pi y_1)\sin(2\pi y_1)(\cos(2\pi y_2)-1)) & \text{on } [-0.8,0.8]^2\setminus[-0.7,0.7]^2, \\ 1, \ \text{else}. \end{cases}$$

Here we are dealing with the bounded domain $[-0.1, 0.1]^2$ instead of $\mathbb{R}^2$ and an additional homogeneous Dirichlet boundary condition. Since b meets the assumptions $\operatorname{div} b(t,\cdot) = 0$ and $\int_Y b(t,\cdot) = 0$, we do not have a macroscopic drift of order $\frac{1}{\epsilon}$. In this case the HMM can be formulated analogously to approximate the corresponding solution. Assuming that the coefficients are periodic, the same a-posteriori and a-priori estimates as in the sections 4 and 5 can be obtained. This example is to focus on the applicability of the HMM in the case of heterogeneous structures within the coefficients, since standard homogenization fails for such problems.

Again, we compare the HMM with a standard Backward-Euler (Linear-)Finite-Element Scheme. The results are essentially the same as for Model Problem 1. In Table 4 we see that the relative HMM error is rapidly decreasing, whereas the BW-FES error remains between the values 0.36 and 0.55 without showing convergence. In Figure 3 we observe that already for a mesh size of $(H, h) = (2^{-4}, 2^{-4})$, the isolines of the exact solution u^ϵ and HMM approximation u_H match very well. In comparison, the isolines of the BW-FES approximation are totally different. Even for the higher refinement level in Figure 4 ($(H, h) = (2^{-5}, 2^{-5})$), the BW-FES approximation has not yet gained a better quality. Instead, the isolines of the HMM approximation are now almost identical with the ones of the exact solution.

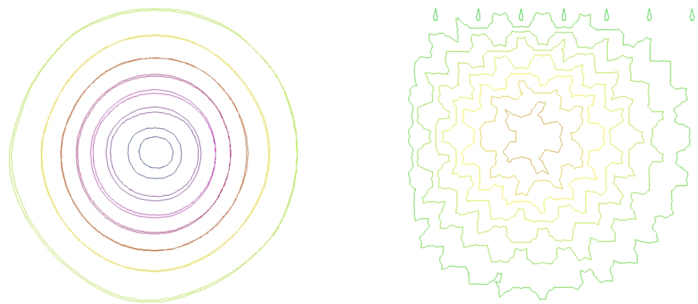

FIGURE 3. *Model problem 2. Here we have* $(H, h) = (2^{-4}, 2^{-4})$ *and* $N = 20$. *The color gradient is from green (minimum* 0*) to blue (maximum* 2.148*). In the left hand figure we see a comparison between the isolines of* $u^{\epsilon}(t^N, \cdot)$ *and* u_H^N. *In the right hand figure the isolines of the BE-FES solution are shown.*

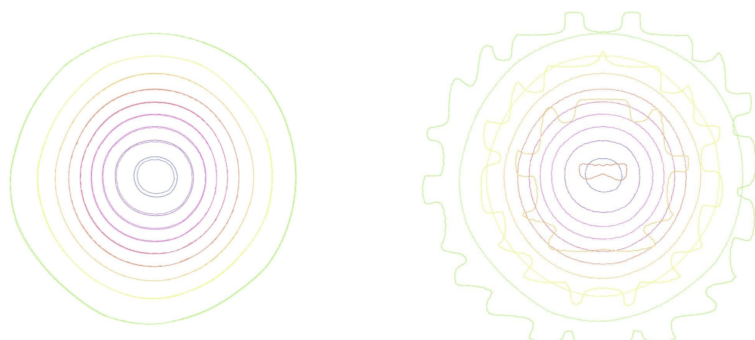

FIGURE 4. *Model problem 2. Here we have* $(H, h) = (2^{-5}, 2^{-5})$, $N = 20$ *and a color gradient from green (minimum* 0*) to blue (maximum* 2.148*). In the left hand figure we again compare the isolines of* $u^{\epsilon}(t^N, \cdot)$ *and* u_H^N. *In the right hand figure we compare the isolines of the exact solution with the ones of the BE-FES solution.*

First correlations between the isolines of BE-FES approximation and the exact solution start to show up at a mesh size of $H = 2^{-6}$. Here, the relative error is $\|e_{BWE}^N\|_{L^2(\Omega)}^{rel} = 0.027$. To obtain a comparable result with the HMM, we can be 4 times coarser.

The experimental orders of convergence in Table 5 show again that the HMM seems to converge with second order in (H, h). The bad results (EOC= 0.69, EOC= 0.87 and to a certain extend EOC= 1.51) are due to the fact, that the time step size $\triangle t$ is too large in comparison to H^2. In Table 6 we observe that the BE-FE Scheme does not converge on coarse grids, which is clear. The time EOC's of the HMM seem to be a little too small, since we expect values around 1. Again, this observation is related to the fact that the mesh size is not yet small enough in comparison to the time step size. For highly refined grids the results will be probably better, showing a linear behaviour. Nevertheless, we note that Model Problem 2 includes a heterogeneous diffusion matrix, which implies that the Theorem 4.1 is

TABLE 5. *Model problem 2. With Definition (25), we calculate the following EOC's in space, i.e.* $EOC_{(2(H,h)\to(H,h))}$*. For each computation* N *is fixed.*

N	$H \to k \cdot H$	$h \to k \cdot h$	EOC(e^N)	EOC(e^N_{BWE})
20	$2^{-\frac{5}{2}} \to 2^{-\frac{7}{2}}$	$2^{-\frac{5}{2}} \to 2^{-\frac{7}{2}}$	2.284	1.1208
20	$2^{-3} \to 2^{-4}$	$2^{-3} \to 2^{-4}$	2.0816	-0.5638
10	$2^{-\frac{7}{2}} \to 2^{-\frac{9}{2}}$	$2^{-\frac{7}{2}} \to 2^{-\frac{9}{2}}$	0.6923	0.3132
20	$2^{-\frac{7}{2}} \to 2^{-\frac{9}{2}}$	$2^{-\frac{7}{2}} \to 2^{-\frac{9}{2}}$	1.5105	0.3002
40	$2^{-\frac{7}{2}} \to 2^{-\frac{9}{2}}$	$2^{-\frac{7}{2}} \to 2^{-\frac{9}{2}}$	2.0603	0.2944
20	$2^{-4} \to 2^{-5}$	$2^{-4} \to 2^{-5}$	0.8758	0.2149

TABLE 6. *Model problem 2. With Definition (25), we calculate the following EOC's in time, i.e.* $EOC_{(k\triangle t\to\triangle t)}$*. For each computation* (H,h) *is fixed.*

$N \to k \cdot N$	H	h	EOC(e^N)	EOC(e^N_{BWE})
$5 \to 10$	$2^{-\frac{9}{2}}$	$2^{-\frac{9}{2}}$	0.8408	-0.1123
$5 \to 10$	2^{-5}	2^{-5}	0.9262	-0.1086
$10 \to 20$	2^{-5}	2^{-5}	0.6728	-0.0515

not applicable. Even though we may expect the same results in most cases, other properties could show up.

General comment: The numerical results have demonstrated the applicability of the HMM of Definition 3.3. The orders of convergence, predicted by Theorem 4.1, could be verified. Results of good quality could be gained with much coarser discretizations of the macro grid than with a comparable Backward-Euler Finite-Element Scheme. Since $\epsilon = 0.01$ was still relative large in comparison to what we could encounter in other problems, this advantage will become much bigger for wider scale separation between micro and macro scale. For several problems, the computational demand for solving the fine-scale equation with a Finite-Element or Finite-Volume-Scheme will be even too high for practical applications. In such cases there is no alternative but a multiscale method. For problems such as Model Problem 1 or 2, one may argue that the computational complexity for solving all the cell problems may be of equal or even larger than a BE-FE Scheme with highly refined grid. But note that all the cell problems are independent from each other, which suggests to solve them in parallel or in a preprocessing step. Assuming that the results of the cell problems are available, the remaining HMM macro problem is only of minor complexity and can be solved very fast.

6. **Conclusion.** In this contribution we formulated the heterogeneous multiscale finite element method for advection-diffusion problems with rapidly oscillating coefficients and large expected drift. For the case of periodic coefficient functions we derived a corresponding a-priori error estimate in the $L^\infty(L^2)$-norm. The convergence is of second order in space and first order in time. In order to demonstrate the applicability and efficiency of the method, numerical experiments were given. One model problem covered the case of a large drift, another model problem the case

of a heterogeneous structure within the diffusion matrix. Even for relatively coarse grids, both problems could be solved with high accuracy. In order to establish a basis for possible adaptive mesh refinement algorithms and error control, we will also apply the techniques of this paper to derive an associated a-posteriori result, based on local error indicators. This will be the subject of future work.

REFERENCES

[1] A. Abdulle, *Multiscale methods for advection-diffusion problems*, Discrete Contin. Dyn. Syst., **suppl** (2005), 11–21.

[2] A. Abdulle, *On a priori error analysis of fully discrete heterogeneous multiscale FEM*, Multiscale Model. Simul., **4** (2005), 447–459 (electronic).

[3] A. Abdulle and W. E, *Finite difference heterogeneous multi-scale method for homogenization problems*, J. Comput. Phys., **191** (2003), 18–39.

[4] A. Abdulle and C. Schwab, *Heterogeneous multiscale FEM for diffusion problems on rough surfaces*, Multiscale Model. Simul., **3** (2004/05), 195–220 (electronic).

[5] G. Allaire and R. Orive, *Homogenization of periodic non self-adjoint problems with large drift and potential*, ESAIM Control Optim. Calc. Var., **13** (2007), 735–749 (electronic).

[6] G. Allaire and A.-L. Raphael, "Homogénéisation d'un Modèle de Convection-Diffusion Avec Chimie/Adsorption en Milieu Poreux," (French), Rapport Interne, CMAP, Ecole Polytechnique, **n. 604**, 2006.

[7] G. Allaire and A.-L. Raphael, *Homogenization of a convection-diffusion model with reaction in a porous medium*, C. R. Math. Acad. Sci. Paris, **344** (2007), 523–528.

[8] T. Arbogast, G. Pencheva, M. F. Wheeler and I. Yotov, *A multiscale mortar mixed finite element method*, Multiscale Model. Simul., **6** (2007), 319–346 (electronic).

[9] A. Bourlioux and A. J. Majda, *An elementary model for the validation of flamelet approximations in non-premixed turbulent combustion*, Combust. Theory Model., **4** (2000), 189–210.

[10] W. E and B. Engquist, *The heterogeneous multiscale methods*, Commun. Math. Sci., **1** (2003), 87–132.

[11] W. E and B. Engquist, *Multiscale modeling and computation*, Notices Amer. Math. Soc., **50** (2003), 1062–1070.

[12] W. E and B. Engquist, *The heterogeneous multi-scale method for homogenization problems*, in "Multiscale Methods in Science And Engineering," Lect. Notes Comput. Sci. Eng., **44**, Springer, Berlin, (2005), 89–110.

[13] W. E, P. Ming and P. Zhang, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., **18** (2005), 121–156 (electronic).

[14] Y. Efendiev and T. Hou, *Multiscale finite element methods for porous media flows and their applications*, Appl. Numer. Math., **57** (2007), 577–596.

[15] V. Gravemeier and W. A. Wall, *A 'divide-and-conquer' spatial and temporal multiscale method for transient convection-diffusion-reaction equations*, Internat. J. Numer. Methods Fluids, **54** (2007), 779–804.

[16] P. Henning and M. Ohlberger, *A-posteriori error estimate for a heterogeneous multiscale finite element method for advection-diffusion problems with rapidly oscillating coefficients and large expected drift*, Preprint, Universität Münster, **N-09/09**, 2009.

[17] P. Henning and M. Ohlberger, *A note on homogenization of advection-diffusion problems with large expected drift*, submitted to: ZAA, Journal for Analysis and its Applications, 2010.

[18] P. Henning and M. Ohlberger, *The heterogeneous multiscale finite element method for elliptic homogenization problems in perforated domains*, Numer. Math., **113** (2009), 601–629.

[19] V. Hoang and C. Schwab, *High-dimensional finite elements for elliptic problems with multiple scales*, Multiscale Model. Simul., **3** (2004/05), 168–194 (electronic).

[20] T. Y. Hou and X.-H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., **134** (1997), 169–189.

[21] T. Y. Hou, X.-H. Wu and C. Zhiqiang, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., **68** (1999), 913–943.

[22] L. Jiang, Y. Efendiev and V. Ginting, *Multiscale methods for parabolic equations with continuum spatial scales*, Discrete Contin. Dyn. Syst. Ser. B, **8** (2007), 833–859 (electronic).

[23] E. Marušić-Paloka and A. L. Piatnitski, *Homogenization of a nonlinear convection-diffusion equation with rapidly oscillating coefficients and strong convection*, J. London Math. Soc. (2), **72** (2005), 391–409 (electronic).

[24] A.-M. Matache, *Sparse two-scale FEM for homogenization problems. Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala)*, J. Sci. Comput., **17** (2002), 659–669.

[25] A.-M. Matache and C. Schwab, *Two-scale FEM for homogenization problems*, M2AN Math. Model. Numer. Anal., **36** (2002), 537–572.

[26] P. Ming and P. Zhang, *Analysis of the heterogeneous multiscale method for parabolic homogenization problems*, Math. Comp., **76** (2007), 153–177 (electronic).

[27] J. Nolen, G. Papanicolaou and O. Pironneau, *A framework for adaptive multiscale methods for elliptic problems*, Multiscale Model. Simul., **7** (2008), 171–196.

[28] J. T. Oden and K. S. Vemaganti, *Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. I. Error estimates and adaptive algorithms*, J. Comput. Phys., **164** (2000), 22–47.

[29] M. Ohlberger, *A posteriori error estimates for the heterogeneous multiscale finite element method for elliptic homogenization problems*, Multiscale Model. Simul., **4** (2005), 88–114 (electronic).

[30] C. Schwab and A.-M. Matache, *Generalized FEM for homogenization problems*, in "Multiscale and Multiresolution Methods," Lect. Notes Comput. Sci. Eng., **20**, Springer, Berlin, (2002), 197–237.

[31] V. Thomée, "Galerkin Finite Element Methods for Parabolic Problems," Springer Series in Computational Mathematics, **25**, Springer-Verlag, Berlin, 1997.

[32] K. S. Vemaganti and J. T. Oden, *Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. II. A computational environment for adaptive modeling of heterogeneous elastic solids*, Comput. Methods Appl. Mech. Engrg., **190** (2001), 46–47.

Appendix A. **Proof of Theorem 3.5.** Here, we want give a detailed proof of Theorem 3.5, which is given in two steps. We start with the following lemma (a reformulation for a fixed time step n):

Lemma A.1. *Suppose that $H \gg \epsilon$ and let U_H, u_H^n and $\mathcal{K}_h^{(n)}(u_H^n)$ be defined in Theorem 3.5. If $u_H^0 = v_H^0$, then we have that for $n \geq 1$, $(u_H^n, \mathcal{K}_h^{(n)}(u_H^n)) \in \tilde{I}_H$ is a solution of*

$$(u_H^{n-1}, \Phi_H)_{L^2(\mathbb{R}^N)} = (u_H^n, \Phi_H)_{L^2(\mathbb{R}^N)} + \triangle t E_H^n \left((u_H^n, \mathcal{K}_h^{(n)}(u_H^n)), (\Phi_H, \phi_h)\right)$$

for all $(\Phi_H, \phi_h) \in I_H$. Note that this equation can be decoupled again into macro and micro-scale part by choosing $\Phi_H = 0$.

Proof. We start with the local part. From Definition 3.3 we have

$$\int_{Y_{j,\epsilon}} A_h^\epsilon(t^n, x) \nabla_x R_j^{(n)}(\Phi_H)(x) \cdot \nabla_x \phi_h(x)\, dx + \int_{Y_{j,\epsilon}} \frac{1}{\epsilon} b_h^\epsilon(t^n, x) \cdot \nabla_x R_j^{(n)}(\Phi_H)(x)\, \phi_h(x)\, dx$$
$$= \int_{Y_{j,\epsilon}} k_h^\epsilon(t^n, x) \;\; \frac{1}{\epsilon} \overline{b}_h(t^n) \cdot \nabla_x \Phi_H(x) \;\; \phi_h(x)\, dx, \qquad \forall \phi_h \in W_h^m(Y_{j,\epsilon}).$$

Since $\nabla \Phi_H$ is a constant on $Y_{j,\epsilon}$, we obtain by using the definition of $\mathcal{K}_j^{(n)}(\Phi_H)$ and the transformation formula for $x_j^\epsilon(y) = x_j + \epsilon y$:

$$\int_Y A_h^\epsilon(t^n, x_j^\epsilon(y)) \left((\nabla_x \mathcal{K}_j^{(n)}(\Phi_H))(x_j^\epsilon(y)) + \nabla_x \Phi_H(x_j)\right) \cdot \nabla_x \phi_h(x_j^\epsilon(y)) \;\; dy$$
$$+ \int_Y \frac{1}{\epsilon} b_h^\epsilon(t^n, x_j^\epsilon(y)) \cdot \left((\nabla_x \mathcal{K}_j^{(n)}(\Phi_H))(x_j^\epsilon(y)) + \nabla_x \Phi_H(x_j)\right) \;\; \phi_h(x_j^\epsilon(y)) \;\; dy$$
$$= \int_Y k_h^\epsilon(t^n, x_j^\epsilon(y)) \;\; \frac{1}{\epsilon} \overline{b}_h(t^n) \cdot \nabla_x \Phi_H(x_j) \;\; \phi_h(x_j^\epsilon(y)) \;\; dy, \qquad \forall \phi_h \in W_h^m(Y_{j,\epsilon}).$$

Since $\nabla_y \mathcal{K}_h^{(n)}(\Phi_H)(x_j, y) = \nabla_x \mathcal{K}_j^{(n)}(\Phi_H)(\epsilon y)$, we have

$$\nabla_x \mathcal{K}_j^{(n)}(\Phi_H)(x_j^\epsilon(y)) = (\nabla_y \mathcal{K}_h^{(n)}(\Phi_H))(x_j, \frac{x_j^\epsilon(y)}{\epsilon}).$$

Defining $\tilde{\phi}_h \in W_h^1(Y)$ by $\tilde{\phi}_h(y) := \frac{1}{\epsilon}\phi_h(\epsilon y))$ (and extending it by periodicity), we get and in analogy to $\mathcal{K}_h^{(n)}(\Phi_H)$:

$$\nabla_x (\phi_h(x_j^\epsilon(y)) = (\nabla_y \tilde{\phi}_h)(\frac{x_j^\epsilon(y)}{\epsilon}).$$

We therefore obtain:

$$\begin{aligned}\int_Y A_h(t^n, \frac{x_j^\epsilon(y)}{\epsilon}) \left((\nabla_y \mathcal{K}_h^{(n)}(\Phi_H))(x_j, \frac{x_j^\epsilon(y)}{\epsilon}) + \nabla_x \Phi_H(x_j) \right) \cdot (\nabla_y \tilde{\phi}_h)(\frac{x_j^\epsilon(y)}{\epsilon})\, dy \\ + \int_Y \frac{1}{\epsilon} b_h(t^n, \frac{x_j^\epsilon(y)}{\epsilon}) \cdot \left((\nabla_y \mathcal{K}_h^{(n)}(\Phi_H))(x_j, \frac{x_j^\epsilon(y)}{\epsilon}) + \nabla_x \Phi_H(x_j) \right)\, \epsilon \tilde{\phi}_h(\frac{x_j^\epsilon(y)}{\epsilon})\, dy \\ = \int_Y k_h(t^n, \frac{x_j^\epsilon(y)}{\epsilon})\, \frac{1}{\epsilon} \bar{b}_h(t^n) \cdot \nabla_x \Phi_H(x_j)\, \epsilon \tilde{\phi}_h(\frac{x_j^\epsilon(y)}{\epsilon})\, dy, \qquad \forall \tilde{\phi}_h \in W_h^1(Y).\end{aligned}$$

Because the equation holds for every $\tilde{\phi}_h \in W_h^1(Y)$, it also holds for $\phi_h(x, \cdot)$ almost everywhere in x and with $\phi_h \in L^2(\mathbb{R}^d, W_h^1(Y))$. Since $\frac{x_j^\epsilon(y)}{\epsilon} = y + \frac{x_j}{\epsilon}$ and since every function is y-periodic, we obtain:

$$\begin{aligned}\int_Y A_h(t^n, y) \left(\nabla_y \mathcal{K}_h^{(n)}(\Phi_H)(x_j, y) + \nabla_x \Phi_H(x_j) \right) \cdot \nabla_y \phi_h(x, y)\, dy \\ + \int_Y b_h(t^n, y) \cdot \left(\nabla_y \mathcal{K}_h^{(n)}(\Phi_H)(x_j, y) + \nabla_x \Phi_H(x_j) \right)\, \phi_h(x, y)\, dy \\ = \int_Y k_h(t^n, y)\, \bar{b}_h(t^n) \cdot \nabla_x \Phi_H(x_j)\, \phi_h(x, y)\, dy, \qquad \forall \phi_h \in L^2(\mathbb{R}^d, W_h^1(Y)).\end{aligned} \tag{26}$$

Analogously we get for the global problem:

$$\begin{aligned}&\fint_{Y_{j,\epsilon}} A_h^\epsilon(t^n, x) \nabla_x R_j^{(n)}(u_H^n)(x) \cdot \nabla_x \Phi_H(x_j)\, dx \\ &+ \fint_{Y_{j,\epsilon}} \frac{1}{\epsilon} \left(k_h^\epsilon(t^n, x) \bar{b}_h^j(t^n) - b_h^\epsilon(t^n, x) \right) \cdot \nabla_x \Phi_H(x_j)\, \overline{R_j^{(n)}}(u_H^n)(x)\, dx \\ &= \fint_Y A_h(t^n, y) \left(\nabla_x u_H^n(x_j) + \nabla_y \mathcal{K}_h^{(n)}(u_H^n)(x_j, y) \right) \cdot \nabla_x \Phi_H(x_j)\, dy \\ &\quad + \fint_Y \frac{1}{\epsilon} \left(k_h(t^n, y) \bar{b}_h(t^n) - b_h(t^n, y) \right) \cdot \nabla_x \Phi_H(x_j)\, (u_H^n(x_j) + \epsilon \mathcal{K}_h^{(n)}(u_H^n)(x_j, y))\, dy \\ &= \fint_Y A_h(t^n, y) \left(\nabla_x u_H^n(x_j) + \nabla_y \mathcal{K}_h^{(n)}(u_H^n)(x_j, y) \right) \cdot \nabla_x \Phi_H(x_j)\, dy \\ &\quad + \fint_Y \left(k_h(t^n, y) \bar{b}_h(t^n) - b_h(t^n, y) \right) \cdot \nabla_x \Phi_H(x_j)\, \mathcal{K}_h^{(n)}(u_H^n)(x_j, y)\, dy,\end{aligned}$$

since $\fint_Y \frac{1}{\epsilon}\left(k_h(t^n,y)\bar{b}_h(t^n) - b_h(t^n,y)\right)\cdot \nabla_x \Phi_H(x_j)\ \ u_H^n(x_j)\ \ dy = 0$. Choosing $\Phi_H = u_H^n$ in (26) and adding this summand to the global equation yields:

$$\begin{aligned}
&A_H^n(u_H^n, \Phi_H)\\
&\quad = \sum_{j\in J} |T_j| \left(\int_Y A_h(t^n,y) \left(\nabla_x u_H^n(x_j) + \nabla_y \mathcal{K}_h^{(n)}(u_H^n)(x_j,y) \right) \cdot \nabla_x \Phi_H(x_j)\ \ dy \right)\\
&\qquad + \sum_{j\in J} |T_j| \left(\int_Y \left(k_h(t^n,y)\bar{b}_h(t^n) - b_h(t^n,y)\right) \cdot \nabla_x \Phi_H(x_j)\ \ \mathcal{K}_h^{(n)}(u_H^n)(x_j,y)\ \ dy \right)\\
&\qquad + \int_{\mathbb{R}^d} \int_Y A_h(t^n,y) \left(\nabla_y \mathcal{K}_h^{(n)}(u_H^n)(x,y) + \nabla_x u_H^n(x) \right) \cdot \nabla_y \phi_h(x,y)\ \ dy\, dx\\
&\qquad + \int_{\mathbb{R}^d} \int_Y b_h(t^n,y) \cdot \left(\nabla_y \mathcal{K}_h^{(n)}(u_H^n)(x,y) + \nabla_x u_H^n(x) \right)\ \ \phi_h(x,y)\ \ dy\, dx\\
&\qquad - \int_{\mathbb{R}^d} \int_Y k_h(t^n,y)\ \ \bar{b}_h(t^n) \cdot \nabla_x u_H^n(x)\ \ \phi_h(x,y)\ \ dy\, dx.
\end{aligned}$$

Since the quadrature is exact for piecewise constant functions we finally obtain:

$$A_H^n(u_H^n, \Phi_H) = E_H^n \left((u_H^n, \mathcal{K}_h^{(n)}(u_H^n)), (\Phi_H, \phi_h) \right), \quad \forall (\Phi_H, \phi_h) \in I_H.$$

□

Now the reformulation can be proved.

Proof of Theorem 3.5. For $(\Phi_H, \phi_h) \in X_H(0, \bar{T})$, we define $(\Phi_H^n, \phi_h^n) \in I_H$ by

$$(\Phi_H^{n+1}, \phi_h^{n+1}) := ((\Phi_H)_+^n, \frac{1}{\triangle t} \int_{t^n}^{t^{n+1}} \phi_h(t,\cdot)\ \ dt).$$

With Lemma A.1 we therefore obtain for all $(\Phi_H, \phi_h) \in X_H(0, \bar{T})$

$$\begin{aligned}
&(u_H^n, (\Phi_H)_+^n)_{L^2(\mathbb{R}^N)}\\
&\quad = (u_H^{n+1}, (\Phi_H)_+^n)_{L^2(\mathbb{R}^N)} + \triangle t E_H^{n+1} \left((u_H^{n+1}, \mathcal{K}_h^{(n+1)}(u_H^{n+1})), ((\Phi_H)_+^n, \phi_h^{n+1}) \right).
\end{aligned}$$

Defining $[u_H]_0 := u_H^1 - u_H^0$, we get by summing up:

$$\sum_{n=0}^{N-1} \left(([u_H]_n, (\Phi_H)_+^n)_{L^2(\mathbb{R}^N)} + \triangle t E_H^{n+1} \left((u_H^{n+1}, \mathcal{K}_h^{(n+1)}(u_H^{n+1})), ((\Phi_H)_+^n, \phi_h^{n+1}) \right) \right) = 0.$$

Since $(\Phi_H)_+^n = \Phi_H(t,\cdot)$ for all $t \in (t^n, t^{n+1}]$ and since $\phi_h^{n+1} = \fint_{t^n}^{t^{n+1}} \phi_h(t,\cdot)\ \ dt$, we obtain

$$\sum_{n=0}^{N-1} ([u_H]_n, (\Phi_H)_+^n)_{L^2(\mathbb{R}^N)} + \sum_{n=0}^{N-1} \int_{t^n}^{t^{n+1}} E_H(t) \left((u_H, \mathcal{K}_h(u_H)), (\Phi_H, \phi_h) \right) = 0.$$

Together with $\partial_t (u_H)_{|(t^n,t^{n+1})} = 0$ we finally have the result. □

Appendix B. Proof of Lemma 4.8.

Proof of Lemma 4.8. We denote $\|\cdot\| := \|\cdot\|_{L^2(\mathbb{R}^d)}$ and $(\cdot,\cdot) := (\cdot,\cdot)_{L^2(\mathbb{R}^d)}$. Using (21) and testing with $\Phi_H = z_H^{N+1}$ and $\Phi_H = z_H^{N+1} - z_H^N$, yields the equations:

$$(z_H^{N+1}, z_H^{N+1} - z_H^N) = \triangle t S_H^N (z_H^{N+1}, z_H^N)$$

and

$$\|z_H^{N+1} - z_H^N\|^2 = \triangle t S_H^N(z_H^{N+1}, z_H^N) - \triangle t S_H^N(z_H^N, z_H^N).$$

Combining this and since $S_H^N(\tilde{z}_H^N, \tilde{z}_H^N) \geq 0$ we obtain:

$$\begin{aligned}\|z_H^{N+1} - z_H^N\|^2 &\leq \triangle t S_H^N(z_H^{N+1}, z_H^N)\\ &= (z_H^{N+1}, z_H^{N+1} - z_H^N)\\ &\leq \|z_H^{N+1} - z_H^N\| \|e^N\|.\end{aligned}$$

In particular this yields

$$(t^{N+1} - t^N)\frac{1}{\triangle t}\|z_H^{N+1} - z_H^N\|^2 \leq \|e^N\|^2. \tag{27}$$

Now suppose that $n \leq N-1$. We test with $\Phi_H = z_H^{n+1} - z_H^n$ to obtain:

$$\begin{aligned}&2\|z_H^{n+1} - z_H^n\|^2 = 2\triangle t S_H^n(z_H^{n+1} - z_H^n, z_H^n)\\ &= \quad -\triangle t S_H^n(z_H^{n+1} - z_H^n, z_H^{n+1} - z_H^n) - \triangle t S_H^n(z_H^n, z_H^n) + \triangle t S_H^n(z_H^{n+1}, z_H^{n+1})\\ &\leq \quad -\triangle t S_H^n(z_H^n, z_H^n) + \triangle t S_H^n(z_H^{n+1}, z_H^{n+1}).\end{aligned}$$

Here we used the symmetry of S_H^n. Multiplying with $(t^{N+1} - t^n)\frac{1}{\triangle t}$ and summing up yields:

$$\begin{aligned}&2\sum_{n=0}^{N-1}(t^{N+1} - t^n)\frac{1}{\triangle t}\|z_H^{n+1} - z_H^n\|^2\\ &\quad \leq \underbrace{-\sum_{n=0}^{N-1}(t^{N+1} - t^n)S_H^n(z_H^n, z_H^n) + \sum_{n=0}^{N-1}(t^{N+1} - t^n)S_H^n(z_H^{n+1}, z_H^{n+1})}_{=:\mathrm{I}}.\end{aligned}$$

In order to estimate I, we use

$$\begin{aligned}&S_H^n(z_H^{n+1}, z_H^{n+1}) - S_H^n(z_H^n, z_H^n)\\ &\quad = \quad S_H^{n+1}(z_H^{n+1}, z_H^{n+1}) - S_H^n(z_H^n, z_H^n) + (S_H^n - S_H^{n+1})(z_H^{n+1}, z_H^{n+1}).\end{aligned} \tag{28}$$

Since A_h is a piecewise constant interpolation of a Lipschitz continuous coefficient function and due to $T_h^n(z_H^n) = z_h^n$, we first get:

$$(S_H^{n+1} - S_H^n)(z_H^{n+1}, z_H^{n+1}) \leq C\triangle t\left(\|z_H^{n+1}\|_{H^1(\mathbb{R}^d)} + \|z_h^{n+1}\|_{L^2(\mathbb{R}^d, H^1(Y))}\right)^2.$$

The Young inequality and Lemma 4.6 therefore yield:

$$\sum_{n=0}^{N-1}(t^{N+1} - t^n)(S_H^{n+1} - S_H^n)(z_H^{n+1}, z_H^{n+1}) \leq C\|e^N\|^2_{L^2(\mathbb{R}^d)}. \tag{29}$$

Now we estimate I by making use of (28):

$$\begin{aligned}
\mathrm{I} \leq & \sum_{n=0}^{N-1}(t^{N+1}-t^n)(S_H^n - S_H^{n+1})(z_H^{n+1}, z_H^{n+1}) \\
& + \sum_{n=0}^{N-1}(t^{N+1}-t^{n+1})S_H^{n+1}(z_H^{n+1}, z_H^{n+1}) - \sum_{n=0}^{N-1}(t^{N+1}-t^n)S_H^n(z_H^n, z_H^n) \\
& + \sum_{n=0}^{N-1} \triangle t S_H^{n+1}(z_H^{n+1}, z_H^{n+1}) \\
= & \sum_{n=0}^{N-1}(t^{N+1}-t^n)(S_H^n - S_H^{n+1})(z_H^{n+1}, z_H^{n+1}) + (t^{N+1}-t^N)S_H^N(z_H^N, z_H^N) \\
& - t^{N+1}S_H^0(z_H^0, z_H^0) + \sum_{n=1}^{N} \triangle t S_H^n(z_H^n, z_H^n).
\end{aligned}$$

(29) and again Lemma 4.6 yield:

$$\mathrm{I} \leq C\|e^N\|_{L^2(\mathbb{R}^d)}^2 + 2\sum_{n=1}^{N} \triangle t S_H^n(z_H^n, z_H^n) \leq C\|e^N\|_{L^2(\mathbb{R}^d)}^2. \tag{30}$$

Together with (27), we finally have

$$\sum_{n=0}^{N}(t^{N+1}-t^n)\frac{1}{\triangle t}\|z_H^{n+1}-z_H^n\|_{L^2(\mathbb{R}^d)}^2 \leq C\|e^N\|_{L^2(\mathbb{R}^d)}^2.$$

and therefore

$$\begin{aligned}
& \sum_{n=0}^{N}\frac{1}{\sqrt{\triangle t}}\|z_H^{n+1}-z_H^n\|_{L^2(\mathbb{R}^d)} \\
\leq & \left(\sum_{n=0}^{N}\frac{1}{t^{N+1}-t^n}\right)^{\frac{1}{2}} \left(\sum_{n=0}^{N}(t^{N+1}-t^n)\frac{1}{\triangle t}\|z_H^{n+1}-z_H^n\|_{L^2(\mathbb{R}^d)}^2\right)^{\frac{1}{2}} \\
\leq & C\left(\int_0^{t^N}\frac{1}{t^{N+1}-t}\,dt\right)^{\frac{1}{2}}\|e^N\|_{L^2(\mathbb{R}^d)} \\
= & C\left(\log\frac{t^{N+1}}{\triangle t}\right)^{\frac{1}{2}}\|e^N\|_{L^2(\mathbb{R}^d)}.
\end{aligned}$$

□

Received March 2010; revised May 2010.

E-mail address: patrick.henning@uni-muenster.de
E-mail address: mario.ohlberger@uni-muenster.de