A-POSTERIORI ERROR ESTIMATE FOR A HETEROGENEOUS MULTISCALE APPROXIMATION OF ADVECTION-DIFFUSION PROBLEMS WITH LARGE EXPECTED DRIFT

PATRICK HENNING
Department of Mathematics, KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden

MARIO OHLBERGER
Institut für Numerische und Angewandte Mathematik
Westfälische Wilhelms-Universität Münster
Einsteinstr. 62, D-48149 Münster, Germany

Abstract. In this contribution we address a-posteriori error estimation in $L^\infty(L^2)$ for a heterogeneous multiscale finite element approximation of time-dependent advection-diffusion problems with rapidly oscillating coefficient functions and with a large expected drift. Based on the error estimate, we derive an algorithm for an adaptive mesh refinement. The estimate and the algorithm are validated in numerical experiments, showing applicability and good results even for heterogeneous microstructures.

1. Introduction. In this work we focus on a-posteriori error estimation and adaptivity for a heterogeneous multiscale finite element method (HMM) applied to advection-diffusion problems with rapidly oscillating coefficient functions and large expected drift, i.e. on equations of the following type:

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

Here, $b^\epsilon$ is divergence-free and $\epsilon$ denotes a very small parameter, which is a characteristic size for the microscale property of the problem. Equations of this type have a lot of applications, especially in hydrology. They can be used for describing reservoir displacement problems or the transport of solutes in ground- and surface water. When the water content takes values that are close to saturation, the diffusion process has only a minor influence and the flow is primarily caused by gravity. In this spirit, the scaling of the advective part with $\epsilon^{-1}$ refers to a large Péclet number as suggested by Bourlioux and Majda [18]. In fluid dynamics, the Péclet number describes the ratio of the advective part to the diffusive part. Therefore, a large number indicates an advection dominated transport process. As mentioned above, this particularly occurs with the transport of solutes in groundwater. Beside this, there is wide range of other applications, including the modeling of semi-conductor devices or polymer chemistry.

One possibility for solving equations of type (1) is related to the technique of homogenization, where the limit problem for $\epsilon \to 0$ is examined. For corresponding results in the periodic setting (i.e. $\epsilon$-periodic coefficients), we refer to contributions

2010 Mathematics Subject Classification. 35K15, 35B27, 65N30, 65M15.

Key words and phrases. Advection-diffusion, HMM, multiscale method, error estimation.
of Allaire and Raphael [15, 16, 38] for advection-diffusion problems with reaction in a porous medium and Allaire and Orive [14] for advection-diffusion-reaction problems with coefficients that vary both on the macro- and the microscale. The case of nonlinear convection-diffusion problems was treated by Marušić-Paloka and Piatnitski [53]. As the range of the very effective homogenization approach is limited to specific structures, such as periodicity, it is more and more important to consider numerical multiscale methods for more general scenarios. These methods can strongly reduce the complexity of the regarded system by decoupling the problem into macroscale and microscale contributions.

A prominent example for multiscale methods is the so called multiscale finite element method developed by Hou et al. The method can be applied to heterogeneous composite materials and also to porous media. Elliptic equations were considered in [41, 42] and two phase flow in porous media were studied in [24]. Multiscale methods for solving parabolic equations with continuum spatial scales and heterogeneous coefficients were discussed in the work of Jiang, Efendiev and Ginting [46]. Other numerical multiscale approaches are the two-scale finite element methods by Schwab and Matache [54, 55, 59, 40], the multiscale mortar mixed finite element discretization by Arbogast et al. [17], or the ‘divide-and-conquer’ spatial and temporal multiscale method for transient advection-diffusion-reaction equations by Gravemeier and Wall [34]. Another very powerful class is the variational multiscale method proposed by Hughes et al. [44, 45]. Here the solution is split into a fine and a coarse-scale part. The corresponding fine-scale equations are solved in dependency of the residual of the coarse-scale solution. Larson and Målqvist derive an a-posteriori estimate via duality techniques for this method. This estimate is used to create an adaptive algorithm. Diffusion dominated elliptic problems are treated in [48, 49] and the case of stationary advection-diffusion problems in [50]. Another general framework for adaptive multiscale methods for elliptic problems is suggested by Nolen, Papanicolaou and Pironneau [57].

In our contribution we focus on the heterogeneous multiscale finite element method (HMM) which is contrary to most other multiscale methods not designed to approximate the exact solution $u^\epsilon$, but to approximate an homogenized solution $u_0$ which contains no fine scale oscillations. The HMM was originally introduced by E and Engquist in 2003 [20, 21, 22]. It is based on a standard finite element approach for the macroscopic part of the problem. For the evaluation of the corresponding discrete bilinear form, the solutions of so called cell problems are used. These solutions, only defined on small cells around quadrature points, help to reconstruct the efficient macroscopic properties of the coefficient functions. The method is not restricted to the periodic case but requires scale separation.

In the works of E, Ming and Zhang [23], Abdulle and Schwab [12], Ohlberger [58] and Henning and Ohlberger [36] the elliptic case is treated. The parabolic case is analysed by Abdulle and E [6], Abdulle and Huber [9], Ming and Zhang [56] and Henning and Ohlberger [37]. An algorithm for solving advection-diffusion problems is suggested by Abdulle [1]. A HMM for the wave equation was proposed and analyzed by Abdulle and Grote [8] and Engquist, Holst and Runborg [26, 27]. A higher-order HMM based on least-square reconstruction was introduced in [51]. A combination of the HMM with a Reduced Basis approach can be found in [4, 5]. For an overview on the HMM we refer to work by Abdulle, E, Engquist and Vanden-Eijnden [7].
Among others, a-priori results concerning HMM were achieved by E et al. [20, 23], Abdulle [1, 2], Ohlberger [58], Henning and Ohlberger [36, 37], Du and Ming [19] and Abdulle and Vilmart [13] and Gloria [31, 32]. First a-posteriori error estimates for the heterogeneous multiscale method for elliptic problems can be found in Ohlberger [58] and Henning and Ohlberger [36] in the periodic setting. For a posteriori error estimates for the HMM in general (non periodic) settings, we refer to [10, 11, 39]. Another approach towards a-posteriori error estimation is given by Larson and Målvqvist [50] in the context of the variational multiscale method (VMM). However, the techniques applied for the VMM do not generalize to the HMM.

The goal of this work is to derive an a-posteriori error estimate for the heterogeneous multiscale method for advection diffusion problems with rapidly oscillating coefficient functions and a large expected drift. This method was introduced in [37] and is constructed to capture the effective global properties of the solution \( u^r \) of problem (1), it is not to determine \( u^r \) itself. The a-posteriori error estimate proposed in this contribution is derived under the assumption of periodicity. More precisely, we establish the estimate by using the result that the method is equivalent to a discretization of the two-scale homogenized equation of (1) by means of a discontinuous-Galerkin time stepping method. The derived estimate contains local error indicators, which allow the use of adaptive mesh refinement algorithms. Although we prove the estimate under the restriction of a periodic micro-structure, the result is also applicable to ergodic stochastic coefficients or moderately heterogeneous structures. This claim is demonstrated by two numerical experiments.

Outline: In Section 2 we state our heterogeneous multiscale method and present our main results: an a-posteriori error estimate, and a strategy for adaptive mesh refinement. In Section 3 we discuss two numerical experiments to validate our claims and the achieved estimate. Section 4 is dedicated to the proof of the a posteriori result.

2. Formulation of the method and a-posteriori error estimate. In the following we study the advection-diffusion problem (1) with rapidly oscillating and time-dependent coefficient functions. In addition to standard assumptions that guarantee existence and uniqueness of (1), such as ellipticity of \( A^t \) uniformly in \( t \) and \( x \), we demand that \( k^r \) is a positive function and that \( b^r \) is divergence-free. If the advective effects, created by \( b^r \) are expected to average out (i.e. \( \int_{x_T \in [0, \epsilon]^d} b^r(t, \cdot) = 0 \) for every relevant quadrature point \( x_T \in \mathbb{R}^d \)), we do not need further assumptions on \( A^t \), except that there is some kind of scale separation. If this is not the case, we need to presume that \( A^t(t, \cdot) \) and \( b^t(t, \cdot) \) are only micro-scale functions, i.e. they only show a microscopic behaviour and are almost constant on the macro-scale. Examples are periodic coefficients or ergodic stochastic coefficients. For \( k^r \) we assume that the function is \( \epsilon \)-periodic with average 1. We note that this assumption is not a real restriction, but a simplification. The case with a completely general \( k^r \) yields no further difficulties. See for instance [37] on how to formulate the subsequent multiscale method for this case.

Before stating the HMM, we introduce the following notations and definitions: \( T_h \) defines a regular simplicial partition of \( \mathbb{R}^d \), the corresponding set of the inner faces is defined by \( \Gamma(T_h) := \{ E \mid E = T \cap \tilde{T} \neq \emptyset, T, \tilde{T} \in T_h \} \) and by \( x_T \) we denote the barycenter of \( T \in T_h \). \( T_h \) defines a regular periodic partition of the zero-centered unit cube \( Y := [-\frac{1}{2}, \frac{1}{2}]^d \). The set of inner faces of \( T_h \) is given by \( \Gamma(T_h) := \{ E_Y \mid E_Y = S \cap \tilde{S} \neq \emptyset, S, \tilde{S} \in T_h \} \). Furthermore, for \( \delta \in \mathbb{R}_{\geq 0} \), we introduce the \( \delta \)-scaled unit-cells, centered around a barycenter \( x_T \) by \( Y_{T, \delta} := \{ x_T + \delta y \mid y \in Y \} \). The associated
bijection $x^\delta_T : Y \rightarrow Y_{T, \delta}$ is given by $x^\delta_T(y) := x_T + \delta y$, for $y \in Y$ and its extension $x^\delta : \mathbb{R}^d \times Y \rightarrow \bigcup_{T \in T_H} Y_{T, \delta}$ by $x^\delta(x, y) := x^\delta_T(y)$, if $x \in T$. In the following, equidistant time steps will be used for simplification. We define $t^n := n \Delta t$, where $\Delta t$ denotes the step size, such that $N := \frac{T}{\Delta t} \in \mathbb{N}$. The jump over $t^n$ is defined by $[u]_n := u^n - u^n_\star$, where $u^n_\star := \lim_{\epsilon \searrow 0} u(t, \cdot)$, $u^n := \lim_{\epsilon \searrow 0} u(t, \cdot)$. Furthermore, for any bounded domain $M$, we denote the outer normal by $\nu_M : \partial M \rightarrow \mathbb{R}^d$. For two domains $M_1$ and $M_2$, with $\Gamma := \overline{M}_1 \cap \overline{M}_2$ and for a function $g \in (L^\infty(M))^d$ with $g_{M_i} \in (C^0(M_i))^d$, $i = 1, 2$, we define the jump $[g]_{\Gamma} : \Gamma \rightarrow \mathbb{R}$ of $g$ over $\Gamma$ by

$$[g]_{\Gamma}(x) := \lim_{m_1 \rightarrow \infty} g(x_{m_1}) \cdot \nu_{M_1}(x) + \lim_{m_2 \rightarrow \infty} g(x_{m_2}) \cdot \nu_{M_2}(x),$$

where $x_{m_i}$ is a sequence in $M_i$, with $x_{m_i} \rightarrow x$.

We introduce a Sobolev space of periodic functions with zero mean by $\tilde{H}^1_L(Y) := \{ v \in H^1(Y) \mid \int_{\Gamma_n} v(y) \, dy = 0, \phi \in Y \text{ periodic} \}$. For the HMM itself, we need the subsequent discrete spaces:

$$V_H := \{ \Phi_H \in H^1(\mathbb{R}^d) \cap C^0(\mathbb{R}^d) \mid \Phi_{H,t} \in \mathbb{P}^1(T) \ \forall T \in T_H \};$$

$$W_h(Y) := \{ \phi_h \in H^1(Y) \cap C^0(Y) \mid \phi_{h,S} \in \mathbb{P}^1(S) \ \forall S \in T_h \};$$

$$W_{h}(Y, \delta) := \{ \phi_h \in H^1(Y, \delta) \mid (\phi_h \circ x^\delta_T) \in W_h(Y) \};$$

$$V^p_H(\mathbb{R}^d, W_h(Y)) := \{ \phi_h \in L^2(\mathbb{R}^d, \tilde{H}^1_L(Y)) \mid \phi_{h}(\cdot, y)_{|T} \in \mathbb{P}^0(T) \ \forall T \in T_H, y \in Y;$$

$$\lim_{m \rightarrow \infty} \phi_h(x, \cdot) \in W_h(Y) \ \forall x \in \mathbb{R}^d \}.$$

Assuming that the coefficient functions are continuous, we use the following discrete approximations of the coefficients $A^e$, $b^e$ and $k^e$ (in $\bigcup_{T \in T_h} Y_{T, \delta}$):

$A^e_h(t, x) := A^e(t^n, x^n_T(y_S))$, $b^e_h(t, x) := b(t^n, x^n_T(y_S))$, $k^e_h(t, x) := k(t^n, x^n_T(y_S))$

For $(t, x) \in [t^n, t^{n+1}] \times x^n_T(S)$, the local mean of the discrete advection $b^e_h$ is defined as $\bar{b}^e_h(t) := \int_{x^n_T(S)} b^e_h(y, t) \, dy$ with some $\delta > 0$ that is specified in the method itself.

Using a Newton-Cotes quadrature formula of order zero, the heterogeneous multiscale finite element method for advection-diffusion problems with rapidly oscillating coefficient functions is introduced in the subsequent definition. Note that the method is only designed to capture the effective macroscopic properties of $u^\epsilon$ but not its fine-scale oscillations. As proved in [37], in the periodic setting there is a clear relation between HMM approximation $u_H$ and homogenized solution $u_0$, in the sense that $u_H$ converges to $u_0$ strongly in $L^2$.

**Definition 2.1.** The HMM approximation $U_H$ of $u^\epsilon$ is defined through

$$U_H(t^n, x) := u^n_H(x - \frac{1}{\epsilon} \int_0^{t^n} \int_{Y_{T,s}} b^e_h(y, s) \, dy \, ds)$$

where $u^{n+1}_H \in V_H$ is the solution of

$$(u^*_H, \Phi_H)_{L^2(\mathbb{R}^N)} = (u^{n+1}_H, \Phi_H)_{L^2(\mathbb{R}^N)} + \Delta t A^e_{H}^{n+1}(u^{n+1}_H, \Phi_H) \quad \forall \Phi_H \in V_H,$$

with the macroscopic bilinear form $A^e_H$ defined as

$$A^e_{H}(u_H, \Phi_H) := \sum_{T \in T_H} |T| \int_{Y_{T,s}} A^e_h(x, t^n) \nabla x R^e_T(u_H)(x) \cdot \nabla x \Phi_H(x_T) \, dx$$

$$+ \sum_{T \in T_H} |T| \int_{Y_{T,s}} \frac{1}{\epsilon} (k^e_h(x, t^n) \bar{b}^e_h(x, t^n) - b^e_h(x, t^n)) \cdot \nabla x \Phi_H(x_T) \bar{R}^e_T(u_H)(x) \, dx.$$
Here, the local centered reconstructions $\overline{R}_T^{(n)}$ are given as
\[ \overline{R}_T^{(n)}(\Phi_H)(x) := R_T^{(n)}(\Phi_H)(x) - (\Phi_H(x) - \Phi_H(x_T)), \]
and $R_T^{(n)}$ denotes the local reconstruction operator defined through the following cell problems. The image $R_T^{(n)}(\Phi_H) \in E_T(\Phi_H) + W_H(Y_{T,\delta})$ of $\Phi_H \in V_H$ under $R_T^{(n)}$ is the solution of the cell problem
\[
\int_{Y_{T,\delta}} A_h^T(\cdot, t^n) \nabla R_T^{(n)}(\Phi_H) \cdot \nabla \phi_h + \int_{Y_{T,\delta}} \epsilon^{-1} b_h^T(\cdot, t^n) \cdot \nabla R_T^{(n)}(\Phi_H) \cdot \phi_h \\
= \int_{Y_{T,\delta}} k_h^T(\cdot, t^n) \epsilon^{-1} b_h^T(t^n) \cdot \nabla \Phi_H \cdot \phi_h \quad \forall \phi_h \in W_H(Y_{T,\delta}),
\]
where $E_T : \mathbb{P}^1(T) \to \mathbb{P}^1(\mathbb{R}^d)$ denotes the canonic extension operator. The initial value $u_H^0 = v_H^0$ is given by a suitable discretization of $v_H$. For the parameter $\delta$ we assume $\delta \geq \epsilon$. An expedient choice for the periodic case is $\delta = \epsilon$. For $\delta > \epsilon$, the strategy is called oversampling and it is used to erase the effects of a possibly wrong boundary condition for the cell problems.

Practically, it is of great interest to have some a priori knowledge on how to choose the oversampling parameter $\delta$. Until now, there is no general answer to this question, since corresponding error estimates fundamentally rely on the homogenization setting that we are in. For instance, consider a linear elliptic homogenization problem with $A(\cdot) = A(\cdot)$ and a corresponding HMM where the local problems are solved in $Y_{T,\delta}$-cells, but with a homogenous Dirichlet boundary condition. For the averaging in the global bilinear form, the $Y_{T,\delta}$-cells are used instead of smaller cells (i.e. no oversampling). In this case, the error produced by an inappropriate choice of $\delta$ is of order $\frac{1}{\delta}$ (c.f. [23, Theorem 1.2] and [3, Theorem 16 and 17]), which means that $\delta$ should be significantly larger than $\epsilon$. If we replace the periodic homogenization setting by a stochastic homogenization setting, the order of the error term can degenerate to $(\frac{1}{\delta})^\kappa$ with some $\kappa \leq 1$ (c.f. [23, Theorem 1.3]). The same techniques lead to similar results for the MsFEM (c.f. [25, 43]). However, practically and using a periodic boundary condition in combination with oversampling, these rates do not become visible, unless $H$ becomes very small. This indicates that the pre-factor in front of the error terms is often small. In order to derive similar estimates as in the aforementioned works ([23] and [3]) for the HMM with large expected drift, the same strategies can be used where the oversampling can be treated as in [43]. However, one has to consider an asymptotic expansion (c.f. [47]) of the solution to a periodic homogenization problem of a stationary advection-diffusion equation with a periodic boundary condition. Deriving bounds for the various terms in such an estimate is significantly complicated by the periodic boundary condition. To our knowledge, an analysis of such cases was not yet carried out.

An alternative to the usage of a periodic boundary condition for the local problems was proposed in [33] for the elliptic case. Here, Dirichlet boundary conditions are used in combination with an additional regularization of the cell problems by adding the term $\kappa^{-1}(R_T(\Phi_H) - \Phi_H, \phi_h)_{L^2(Y_{T,\delta})}$ (for large positive $\kappa$) to the left hand side of the cell problems. The new problem is analyzed using Green functions. This modified approach improves the estimates in the periodic and stochastic setting considerably.

However, no results for general cases are available. The reason why it is extremely difficult to generalize results from periodic (or stochastic) settings to other types of...
problems is related to the fact that the HMM is always constructed to approximate the homogenized solution, but not to approximate the exact solution. If we are in a setting where do not know the homogenized solution (or how to characterize it) there do not yet exist any tools that allow to derive explicit convergence rates, which can be used for a more general numerical analysis of the HMM. Exact methods like the VMM do not suffer from this since they approximate the exact solution. Hence, more precise (and more general) answers to the question of how big sampling domains should be are possible (c.f. [52]). Practically, the HMM has an advantage in cases, where there is no need for resolving the microstructure everywhere, or where data is only available in small representative volume elements (which would be the $Y_{T,\delta}$ cells).

We also note that the choice of $\delta$ has no influence on the approximation of the speed of the drift. The direction of the drift on the other hand is purely determined by a local average $\delta^{-d} \int_{Y_{T,\delta}} b'$. Assume that we are in the periodic setting, where the exact period $\epsilon$ is unknown. In this case, $\epsilon^{-d} \int_{Y_{T,\epsilon}} b'$ is the correct drift value that should be approximated. We easily see that the error is determined by how often the $\epsilon$-cell fits into the $\delta$-cell and an $L^1$-bound for $b'$ on the ‘remainder cell’ divided by the number of ‘$\epsilon$-samples’, i.e. the error behaves like $((\frac{\delta}{\epsilon})^d - (\frac{\delta}{\epsilon})^d) / (\frac{\delta}{\epsilon})^d \cdot \|b'\|_{L^\infty}$. This is an error that is of smaller order then the $\frac{\delta}{\epsilon}$-terms that show up in the mentioned analysis of the homogenization error.

2.1. Homogenization. Our goal is to derive an a-posteriori result to estimate the $L^\infty(L^2)$-error between the HMM approximation and the solution of the two scale homogenized equation of (1). To this end, let us pose the following periodicity and regularity assumption:

**Assumption 2.2.** We suppose that the initial value belongs to $H^1(\mathbb{R}^d)$ and that the coefficients are Lipschitz-continuous and space periodic with period $\epsilon$, i.e. denoting $A(t, y) := A^\epsilon(t, \epsilon y)$, $b(t, y) := b^\epsilon(t, \epsilon y)$ and $k(t, y) := k^\epsilon(t, \epsilon y)$, we assume

$$A(t, y) \in L^1(0, T_0); H^1(\mathbb{R}^d) \quad \forall y \in \mathbb{R}^d. \ a.e. \ in \ (0, T_0) \times Y;$$

$$b(t, y) \in L^2(0, T_0); H^1(\mathbb{R}^d) \quad \forall y \in \mathbb{R}^d. \ a.e. \ in \ (0, T_0); \ v_0 \in H^1(\mathbb{R}^d);$$

$$k(t, y) \in H^1(0, T_0; H^1(\mathbb{R}^d)); \ k > 0; \ \int_Y k(t, y) dt = 1 \ everywhere \ in \ [0, T_0].$$

If $|\cdot|_{H^k(Y)}$ denotes the semi-norm on $H^k(Y)$ and if $\|\cdot\|_{H^k(Y)}$ denotes the full norm, we introduce the subsequent semi-norm and norm on $L^2(\mathbb{R}^d, H^k(Y))$

$$|\Phi|_{L^2(\mathbb{R}^d, H^k(Y))} := \left( \int_{\mathbb{R}^d} \Phi(x, \cdot)^2_{H^k(Y)} dx \right)^{\frac{1}{2}}; \ \|\Phi\|_{L^2(\mathbb{R}^d, H^k(Y))} := \sum_{l=0}^{k} |\Phi|_{L^2(\mathbb{R}^d, H^k(Y))}.$$

In order to define the two-scale homogenized equation, we introduce the solution space on $[0, t] \subset [0, T_0]$ by

$$X(0, t) := H^1(0, t; H^1(\mathbb{R}^d)) \times L^2(0, t) \times \mathbb{R}^d, H^1_{\|\cdot\|}(Y).$$

With the averaged advection field defined as $\bar{b}(t) := \int_Y b(t, y) dy$ we introduce the elliptic part of the two-scale homogenized operator $E'$ at $t \in [0, T_0]$ by
A-posteriori error estimate


de(x, y) = \sum_{n=0}^{N-1} \left( \int_{t^n}^{t^{n+1}} (\partial_t u_0, \Phi)_{L^2(\mathbb{R}^d)} + E(t)((u_0, u_1), (\Phi, \phi)) \right)
+ \sum_{n=1}^{N-1} \left( [u_0]_n, \Phi^0_n \right)_{L^2(\mathbb{R}^d)} + \left( (u_0)^0_n, \Phi^0_n \right)_{L^2(\mathbb{R}^d)}.

(3)

With these definitions, we are prepared to formulate the following two-scale homogenization result (cf. [38] and the references therein).

**Theorem 2.3.** Let \( u^\varepsilon \) denote the solution of (1) and suppose that the assumptions of this section hold true. Then there exists a unique solution \((u_0, u_1) \in X(0, T_0)\) of the two-scale homogenized equation with drift:

\[
G^N((u_0, u_1), (\Phi, \phi)) = (v_0, \Phi^0_+))_{L^2(\mathbb{R}^d)}
\]

(4)

for all \((\Phi, \phi) \in L^2(0, T_0; H^1(\mathbb{R}^d) \times L^2(0, T_0) \times \mathbb{R}^d, H^2_0(Y))\) with piecewise continuous \(\Phi(\cdot, x)\) (i.e., \(\Phi|_{[t^n, t^{n+1}]} \in C^0([t^n, t^{n+1}; H^1(\mathbb{R}^d)])\)).

Moreover, \(u^\varepsilon\) converges towards \(u_0\) in the following sense

\[
\|u_0(t, \cdot) - \varepsilon^{-1} \int_0^t \bar{b}(s) \, ds - u^\varepsilon(t, \cdot)\|_{L^2(\mathbb{R}^d)} \to 0 \text{ for } \varepsilon \to 0
\]

(5)

and the subsequent regularity for the homogenized solution holds:

\[
\begin{align*}
  u_0 & \in H^1(0, T_0; H^1(\mathbb{R}^d)) \cap L^2(0, T_0; H^2(\mathbb{R}^d)) \text{ and } \\
  u_1 & \in L^2(0, T_0; H^1(\mathbb{R}^d, H^2_0(Y))) \cap L^2((0, T_0) \times \mathbb{R}^d, H^2(Y)).
\end{align*}
\]

Furthermore, we have the estimates

\[
|u_1|_{L^2(0, T_0) \times H^2(\mathbb{R}^d)} \leq C |u_0|_{L^2(0, T_0; H^1(\mathbb{R}^d)} \leq C |v_0|_{L^2(\mathbb{R}^d)}.
\]

(6)

### 2.2. A posteriori error estimate

In the following, the notation \( f \lesssim g \) is used if \( f \leq Cg \), where \( C \) only depends on the domain, the coefficients and the initial value of our problem. We are now prepared to state our main result.

**Theorem 2.4.** (A-posteriori error estimate) Let \( u_H^N, u_0^N \) denote the HMM solution from Definition 2.1, and the macroscopic part of the two-scale homogenized equation (4) for \( t = t^N \) and \( \delta = \varepsilon \), respectively. Under Assumption 2.2 and the general assumptions of this section, the following a-posteriori error estimate holds true:

\[
\|u_0(t^N, \cdot) - u_H^N\|_{L^2(\mathbb{R}^d)} \lesssim \eta_H^N + \xi_H^N.
\]

(7)

Here, \( \eta_H^N := (\sum_{T \in T_H} (\eta_H^N(T))^2)^{\frac{1}{2}} \) and \( \xi_H^N := (\sum_{T \in T_H} \sum_{n=0}^{N-1} (\xi_H^n(T))^2)^{\frac{1}{2}} \) denote the residual error indicator and the approximation error indicator, respectively. Using
the notation $R^{(n)}(\Phi_H)(x) := R^{(n)}_{\tau}^{(n)}(\Phi_H)(x)$ for $\Phi_H \in V_H$ and $x \in T$, the local constituents of the residual and approximation error indicators are given as follows.

With $K^{(n)}(u_H^n) := R^{(n)}(u_H^n) - u_H^n$, $d_h := (k_h^f b_h^f - b_h^r)$ and $d := (k^f - b^r)$ the local approximation error indicator reads

$$
\xi_H^n(T) := \left( \sum_{S \in T_h} |T| \left( (A_h^n(t^n, \cdot) - A^n(t, \cdot)) \nabla_x R^{(n)}_T(u_H^n) \right) \cdot x_T^n \right) \|x_T^n\|_{L^2((t^n,t^{n+1}) \times S)}^{\frac{1}{2}}
+ \left( \sum_{S \in T_h} |T| \left( (d_h(t^n, \cdot) - d(t, \cdot)) K_T^{(n)}(u_H^n) \right) \cdot x_T^n \right) \|x_T^n\|_{L^2((t^n,t^{n+1}) \times S)}^{\frac{1}{2}}
+ \left( \sum_{S \in T_h} |T| \left( (d_h(t^n, \cdot) - d(t, \cdot)) \nabla_x K_T^{(n)}(u_H^n) \right) \cdot x_T^n \right) \|x_T^n\|_{L^2((t^n,t^{n+1}) \times S)}^{\frac{1}{2}}
+ \left( \sum_{S \in T_h} |T| \left( (b_h(t^n, \cdot) - b(t, \cdot)) \nabla_x K_T^{(n)}(u_H^n) \right) \cdot x_T^n \right) \|x_T^n\|_{L^2((t^n,t^{n+1}) \times S)}^{\frac{1}{2}}.
$$

The local residual error indicator splits into time and space constituents:

$$
\eta_H^n(T) := \eta_{\text{space}}^N(T) + \eta_{\text{time}}^N(T),
$$

where

$$
\eta_{\text{time}}^N(T) := (\Delta t + H^2) \left[ \frac{1}{2} \left( \frac{b_H - u_H}{\Delta t} \right) \|b_H - u_H\|_{L^2(T)}^2 \log \left( 1 + \frac{1}{\Delta t} \right) \right]^{\frac{1}{2}}
+ \left( \sum_{E \in \Gamma(T_H) \setminus \Gamma(T_H) \setminus \cdot \in T \setminus \emptyset} \frac{1}{2} \left( \frac{1}{2} H^{E,2} \eta_{\text{max}}^{(2)} \right) \left( \frac{t_N}{\Delta t} \right) \right)^{\frac{1}{2}}
+ \left( \sum_{E \in \Gamma(T_H) \setminus \Gamma(T_H) \setminus \cdot \in T \setminus \emptyset} \frac{1}{2} \left( \frac{1}{2} H^{E,2} \eta_{\text{max}}^{(2)} \right) \left( \frac{t_N}{\Delta t} \right) \right)^{\frac{1}{2}}
+ \left( \sum_{E \in \Gamma(T_H) \setminus \Gamma(T_H) \setminus \cdot \in T \setminus \emptyset} \frac{1}{2} \left( \frac{1}{2} H^{E,5} \eta_{\text{max}}^{(5)} \right) \left( \frac{t_N}{\Delta t} \right) \right)^{\frac{1}{2}}
$$

and

$$
\eta_{\text{space}}^N(T) := \left( \sum_{n=1}^N \Delta t \sum_{E_Y \in \Gamma(T_h)} h_{E_Y}^3 \mu_{n,T,E_Y}^{(1)} \right)^{\frac{1}{2}}
+ \left( \sum_{n=1}^N \Delta t \sum_{S \in \Gamma(T_h)} h_S^3 \left( \mu_{n,T,S}^{(3)} + \mu_{n,T,S}^{(4)} \right) \right)^{\frac{1}{2}},
$$

for $\eta_{\text{space},h}^N(T)$ has specific contributions from the macro scale ($H$) and the micro scale ($h$):

$$
\eta_{\text{space},h}^N(T) := \left( \sum_{n=1}^N \Delta t \sum_{E_Y \in \Gamma(T_h)} h_{E_Y}^3 \mu_{n,T,E_Y}^{(1)} \right)^{\frac{1}{2}}
+ \left( \sum_{n=1}^N \Delta t \sum_{S \in \Gamma(T_h)} h_S^3 \left( \mu_{n,T,S}^{(3)} + \mu_{n,T,S}^{(4)} \right) \right)^{\frac{1}{2}},
$$

for $\eta_{\text{space},h}^N(T)$ has specific contributions from the macro scale ($H$) and the micro scale ($h$):
where we use for $1 \leq n \leq N$, $T \in T_H$, $S \in T_h$, $E \in \Gamma(T_H)$ and $E_Y \in \Gamma(T_h)$ the following notation for local contributions of the residual error
\[
\mu^{(1)}_{n,T,E_Y} := \left\| \left( \left( A_h^{(n)}(t^n, \cdot) \nabla_x R^{(n)}(u_H^{(n)}) \right) \circ x' \right)_{E_Y} \right\|_{L^2(T \times E_Y)}^2,
\]
\[
\mu^{(2)}_{n,E,Y} := \int_T \left\| \left( \left( A_h^{(n)}(t^n, \cdot) \nabla_x R^{(n)}(u_H^{(n)}) \right) \circ x' \right)_{E_Y} \right\|_{L^2(E)}^2,
\]
\[
\mu^{(3)}_{n,S} := \left\| \left( b_h^{(n)}(t^n, \cdot) \cdot \nabla_x K^{(n)}(u_H^{(n)}) \right) \circ x' \right\|_{L^2(T \times S)}^2,
\]
\[
\mu^{(4)}_{n,T,S} := \left\| \left( b_h^{(n)}(t^n, \cdot) - K_h^{(n)}(t^n) \right) \cdot \nabla_x u_H^{(n)} \circ x' \right\|_{L^2(T \times S)}^2,
\]
\[
\mu^{(5)}_{n,E,Y} := \int_T \left\| \left( \frac{1}{\epsilon} K^{(n)}(u_H^{(n)})(k_h^{(n)}(t^n, \cdot) \nabla_x u_H^{(n)} - b_h^{(n)}(t^n, \cdot)) \right) \circ x' \right\|_{L^2(E)}^2.
\]
Furthermore, $n_{\text{max}}, \tilde{n}_{\text{max}}$ and $\bar{n}_{\text{max}}$ are defined as
\[
n_{\text{max}} := \arg\max_{1 \leq n \leq N-2} \left( \sum_{T \in T_H} (\Delta t^2 + H^2) \right) \frac{\|u^{n+1} - u^n\|^2}{\Delta t^2}
\]
\[
\tilde{n}_{\text{max}} := \arg\max_{1 \leq n \leq N-1} \left( \sum_{E \in \Gamma(T_H)} H^2 E \mu^{(2)}_{n,E,Y} \right),
\]
\[
\bar{n}_{\text{max}} := \arg\max_{1 \leq n \leq N-1} \left( \sum_{E \in \Gamma(T_H)} H^2 E \mu^{(5)}_{n,E,Y} \right).
\]

The indicators depending on $\mu^{(2)}_{n,E,Y}$ and $\mu^{(5)}_{n,E,Y}$ account for jump residuals on the macro-grid and the indicators depending on $\mu^{(1)}_{n,T,E_Y}$ account for jump residuals on the micro grid. The gradient jumps typically converge like $H^{1/2}$ and $h^{1/2}$ respectively, such that all residual terms are expected to converge with the optimal rate (i.e. $O(H^2 + h^2)$).

The approximation indicators reflect the typical approximation properties of the scheme. The temporal order is slightly polluted by the logarithmic term $(\log \frac{N^2}{\Delta t})^2$, which is a typical artifact from the analysis. For comparable results for standard parabolic equations, we refer e.g. to the work of Eriksson, Johnson and Larsson [29].

Even though the estimate is derived under the restriction of space-periodic coefficients, the final result can also be applied to a more general setting. This claim is emphasized by the numerical experiments in Section 3.

**Remark 2.5.** Assuming that we have the relations $H^2 \lesssim \Delta t$ and $H \lesssim h$, it can be shown, that we have efficiency for the error estimator $\eta^N_H$ in Theorem 2.4. In particular, defining $L_N := \left( \log \frac{N^{N+1}}{\Delta t} \right)^{1/2}$ and
\[
E_{0n} := H^2 \max_{1 \leq n \leq N} \left( \|u_0\|_{L^2([t^n-\Delta t, t^n]; H^2(\mathbb{R}^d))} \right) + \Delta t \max_{1 \leq n \leq N} \left( \|\partial_t u_0\|_{L^2([t^n-\Delta t, t^n] \times \mathbb{R}^d)} \right),
\]
[37] proved that the $L^\infty(L^2)$-error converges with $\|u_0(t^n, \cdot) - u_H^N\|_{L^2(\mathbb{R}^d)} \lesssim E_{0n} L_N + e^{approx}_N$, where $e^{approx}_N$ denotes a data approximation error. On the other hand, after removing the contributions of the coefficient functions in $\eta^N_H$, we get that the remaining parts of $\eta^N_H$ can be estimated in analogy to the proof of Theorem 3.3. in [28]. In this case, we obtain $\eta^N_H \lesssim E_{0n} L_N + \text{data-approximation}$. This yields that (except of data approximation) the estimator converges with the same speed as the
error itself, which gives us the desired global efficiency of the local indicator $\eta^N_H$. Local efficiency on the other hand seems to require more sophisticated tools.

2.3. **Space adaptive algorithm.** In the following we introduce a strategy for adaptive mesh refinement using the local error indicators from Theorem 2.4. This strategy is used in Section 3 in the second numerical experiment to improve error and computational time with respect to computations on uniformly refined grids. In the algorithm we determine the average of the local indicators $\Xi^N_H := \eta^N_H + \xi^N_H$. If the ratio between $\Xi^N_H(T)$ and $\Xi^average_H$ is larger than a given threshold $\sigma_{TOL}$, we mark $T$ for refinement. The algorithm stops when the global estimated error $\Xi^N := \eta^N_H + \xi^N_H$ falls below a given tolerance $TOL$. Concerning the cell problems, we use a uniform micro-mesh $T_h$ with a fixed grid size of $h = 2^{-4}$ for simplicity. The micro mesh is fine enough (in comparison to the resolution of the macro-grid) such that we do not see the influence of the discretization error in the cell problems.

We denote for the $n$th time step $V^n_H := \{ \Phi_H \in H^1(\mathbb{R}^d) \cap C^0(\mathbb{R}^d) \mid (\Phi_H)|_T \in P^1(T) \forall T \in T^n_H \}$ where $T^n_H$ defines the corresponding $n$th triangulation. Since $T^{n+1}_H$ is an adaptively altered version of $T^n_H$ and since we do not allow mesh coarsening but only mesh refinement, we always have $u^n_H \in V^n_H$ for $m \geq n$. This implies that the error indicators $\eta^N_H$ and $\xi^N_H$ are well defined when replacing $T^n_H$ by the finest mesh $T^N_H$. For $N = 1$ we replace $\log(1 - \frac{\xi^N}{\eta^N})$ by 1 in $\eta^N_H$. The algorithm for mesh refinement is summarized in Algorithm 1. The algorithm takes as input parameters $TOL, \sigma_{TOL}, T^{N-1}_H(\mathbb{R}^d), T_h(Y)$ and $u^n_H$ for the preceding time steps $0 \leq n \leq N - 1$.

There are several other possibilities for adaptive algorithms using the local error indicators. Depending on the considered problem we might use a splitting of residual error and approximation error to improve the performance. Furthermore, the

---

**Algorithm 1: adaptiveRefine( TOL, \sigma_{TOL}, T^{N-1}_H(\mathbb{R}^d), T_h(Y) )**

Set $T^N_H(\mathbb{R}^d) := T^{N-1}_H(\mathbb{R}^d)$.
Compute $u^N_H$ with $T^N_H(\mathbb{R}^d)$ and $T_h(Y)$.
Compute $\Xi^N_H := \eta^N_H + \xi^N_H$.

while $\Xi^N_H \cdot \left( \sum_{n=0}^{N} \triangle t \| u^n_H \|^2_{L^2(\mathbb{R}^d)} \right)^{-\frac{1}{2}} > \text{TOL}$ do
  foreach $T \in T^N_H(\mathbb{R}^d)$ do
    compute the local error indicators $\Xi^N_H(T) := \eta^N_H(T) + \sum_{n=0}^{N-1} \xi^n_H(T)$
    Compute the average indicator $\Xi^average_H := \frac{1}{\mu^N_H(\mathbb{R}^d)} (\sum_{T \in T^N_H(\mathbb{R}^d)} \Xi^N_H(T))$.
  end
  foreach $T \in T^N_H(\mathbb{R}^d)$ do
    if $\Xi^N_H(T) \geq \sigma_{TOL} \cdot \Xi^average_H$ then
      mark $T$ for one refinement.
    end
  end
  Refine grid (update $T^N_H(\mathbb{R}^d)$).
  Compute $u^N_H$ with $T^N_H(\mathbb{R}^d)$ and $T_h(Y)$.
  Compute $\Xi^N_H := \eta^N_H + \xi^N_H$.
end
macro-mesh error indicators \( \eta_N^H(T) \) and \( \xi_N^H(T) \) can be used to determine suitable values for the grid size of the micro-mesh \( T_h \). We refer to [58] for such a strategy and corresponding numerical results. Going even further, each \( \eta_N^H(T) \) can be decomposed into constituents depending on \( S \in T_h \). With such a decomposition we can also adaptively refine the computational micro-grid while coupling the resolution with the macro-grid. This involves a balancing between the various local error indicators (macro and micro grid).

We also note that the above algorithm does not involve a time step control. In particular, if the given time step size is too large, the algorithm might fail to fall below a desired tolerance. One way to overcome this problem is to couple the time step size with the average macro mesh size (average over the local mesh sizes \( H_T \)). Another way is to only decrease the time step size if the ratio between two estimated errors (two cycles of the algorithm) is close to 1. These are easy ways to prevent the algorithm from not terminating, but they are of course not optimal. A more detailed analysis concerning the spatial and temporal contributions to the error indicators would probably allow for a more efficient space-time adaptivity, but such analysis is beyond the scope of this paper and is left for further studies.

**Remark 2.6.** Practically it can be difficult to make a proper choice of \( \delta \) if \( \epsilon \) is unknown. Since we do not know the correct boundary condition for the local problems on the \( Y_{T,\delta} \) cells, the only strategy seems to be that we make the \( Y_{T,\delta} \)-cell sufficiently large so that it contains enough ‘optimal sample cells’. The strategy can be described as follows: first, we solve a local problem in a cell of size \( Y_{T,\delta} \) and compute (for this element \( T \)) the corresponding effective global coefficients as they appear in the macroscopic bilinear form \( A_H^\alpha \). Then we increase the size of the \( Y_{T,\delta} \)-cell and repeat the computation. If the new effective global coefficient is close to the old one, the first choice for \( \delta \) was probably a good choice. If there was a discrepancy between the computed values, we do not yet have enough ‘optimal sample cells’ in the chosen \( Y_{T,\delta} \)-cell. Repeating this procedure iteratively, we can stop whenever the determined effective coefficients start to stagnate. In this case we observe the convergence to the correct homogenized coefficients and hence we have a good choice for \( \delta \). This strategy is independent of the derived a posteriori error estimate stated in Theorem 2.4, in the sense that it is not resembled by one of the error terms. However, this strategy can be easily combined with the adaptive algorithm.

The availability of a-posteriori error estimates for the HMM in general non-periodic homogenization settings (such as derived in [39, 10, 11]) incorporate the error made by the choice of \( \delta \) indirectly in a ‘modeling error contribution’. The modeling error essentially measures the error between the real homogenized matrix and the effective coefficient that is computed by the HMM for \( H \to 0 \). However, this modeling error cannot be computed, unless periodicity is assumed. Hence, it cannot be used to adaptively control the size of \( \delta \).

The fully rigorous a-posterior estimates obtained for the VMM (as in [48, 49, 50]) are based on the fact, that the local problems decay to zero outside of the support of coarse basis functions. This completely justifies a homogenous Dirichlet boundary condition for the local problems and opens the door to various techniques. In e.g. [49] it is used that the proper size of a patch can be measured by the normal flow of the local solution over the boundary of the patch. Since the local solutions decay to zero (with up to exponential speed), the normal flow must tend to zero. This gives
a very reliable indicator. However, this and similar strategies do not generalize to the HMM, since there is no decay to zero of the local solutions.

3. **Numerical experiments.** Next, we analyze the HMM and the corresponding error indicators from Theorem 2.4 numerically in a periodic and heterogeneous setting. In particular, we apply the a-posteriori estimate for adaptive mesh refinement. Note that the formal assumptions of the theorem are not fulfilled in all our examples, but only the weak assumptions of Section 2. Nevertheless, the results suggest that the method and the achieved global and local error indicators work fine even with stochastic perturbations of periodic structures and moderately heterogeneous coefficient functions. For a-priori convergence results in a completely periodic setting, we refer to [37].

Formally, we are concerned with problems on unbounded domains. In the implementation, however, we restrict to a bounded computational domain. In general, this may produce a globally wrong behaviour of the solution, such as boundary layers or reflection. The most reliable strategy to (almost) avoid these effects is the usage of absorbing boundary conditions. For advection-diffusion problems such boundary conditions are derived by Halpern [35] or Nataf, Rogier and de Sturler [30]. In our test problems, however, it was sufficient to use homogeneous Dirichlet boundary condition for the HMM macro problem on a sufficiently large computational domain.

In both model problems, we make use of the subsequent notations. By $u$ we denote the reference solution of the regarded problem, i.e. the function with which we compare our HMM approximation, to obtain an accurate value for the $L^2$-error. In the first model problem $u$ is the homogenized solution $u_0$ (cf. Section 2.1) and in the second model problem $u$ is $u'(t, x + \frac{R(t)}{\varepsilon})$ (a shift of the exact solution). In order to reduce the effects of a possibly wrong periodic boundary condition for the cell problems in the heterogeneous multiscale method, we use an oversampling technique by choosing $\delta = 2\varepsilon$ for all numerical tests (cf. Definition 2.1).

To distinguish between computations on uniformly and adaptively refined grids, we denote the corresponding numerical solutions for the $n$'th time step by $u_{H}^n$ and $u_{A}^n$, respectively. The overall estimated error for one time step is denoted as $\Xi^N_H := \eta^N_H + \xi^N_H$. The relative errors are defined by $\|e^N_u\|_{L^2(R^2)}^{rel} := \frac{\|u(t^n, \cdot) - u_{H}^n, \cdot\|_{L^2(R^2)}}{\|u(t^n, \cdot)\|_{L^2(R^2)}}$ and $\|e^N_a\|_{L^2(R^2)}^{rel}$ analogously. The relative estimated error is given by $\Xi^N_H^{rel} := \frac{\Xi^N_H}{\|u(t^n, \cdot)\|_{L^2(R^2)}}$ and the relative residual error indicator $\eta^N_H^{rel}$ in the same way. Comparing a certain uniform computation with a corresponding adaptive computation, we introduce:

$$\vartheta_e := \frac{\|e^N_u\|_{L^2(R^2)} - \|e^N_a\|_{L^2(R^2)}^{rel}}{\|e^N_u\|_{L^2(R^2)}},$$ the relative improvement in error, \hfill (8)

$$\vartheta_t := \frac{t_u^{CPU} - t_a^{CPU}}{t_u^{CPU}},$$ the relative improvement in computational time. \hfill (9)

Here $t_u^{CPU}$ and $t_a^{CPU}$ denote the CPU times for the uniform and the adaptive computation. We note that $t_a^{CPU}$ also contains the time that is required for computing the estimated error and the local error indicators, whereas $t_u^{CPU}$ only contains the time for assembling and solving the HMM operator.
Figure 1. Model problem 1, variance $\sigma = 0.1$. Visualization of a zoomed part of the diffusion coefficient function $a_{ij}(x)$ (left) and comparison of the isolines for $u^0$ and for $u^N_{H,u}$ at $t = 3$ (right). The mesh sizes are $H = 2^{-9}$, $h = 2^{-4}$.

For $m \in \mathbb{N}^+$, $\delta \in \mathbb{R}_+^m$ and the error function $g : \mathbb{R}_+^m \to \mathbb{R}_+$ we define the experimental order of convergence (EOC) of $g$ in $(2\delta \to \delta)$ by

$$
\text{EOC}_{(2\delta \to \delta)}(g) := \frac{\log \left( \frac{g(2\delta)}{g(\delta)} \right)}{\log(2)}.
$$

For simplicity, we do not refer to $m$ and $\delta$ in our results. Instead we refer to the numbers of corresponding uniform computations.

**Model problem 1.** Find $u^\epsilon \in L^2((0, 3; H^1(\mathbb{R}^2)))$, with

$$
\partial_t u^\epsilon - \epsilon \nabla \cdot (A'(x) \nabla u^\epsilon) + b'(x) \cdot \nabla u^\epsilon = 0 \quad \text{in } (0, 3) \times \mathbb{R}^2
$$

and

$$
u^0(0, x_1, x_2) = \begin{cases} 
\sin(2\pi x_1) \sin(2\pi x_2) & \text{in } [0, 0.5]^2 \\
2\cos(5\pi x_1) \cos(5\pi x_2) & \text{in } [-0.3, -0.1]^2 \\
0, \quad \text{else.}
\end{cases}
$$

We choose $\epsilon = 0.001$. The entries of $A^\epsilon$ are given by

$$
a_{ij}'(x) = 10 \delta_{ij} \left( 1 + \sin^2(2\pi x_1 / \epsilon) \cos^2(2\pi x_2 / \epsilon) + X(x) - E(X) \right),
$$

where $X$ is a log-normal distributed random variable, with variance $\sigma$ and expectation $E(X) = e^{\frac{\sigma^2}{2}}$ (see Fig. 1 for $a_{ij}'$ with $\sigma = 0.1$). If $a_{ij}'(x)$ would take a value below $\epsilon$, we set it to $\epsilon$ to keep the ellipticity. The advection $b'$ is defined by

$$
b'(x) = \left( \frac{30}{\sqrt{2}} \sin(2\pi x_1 / \epsilon) \sin(2\pi x_2 / \epsilon) + 20, \quad \frac{30}{\sqrt{2}} \cos(2\pi x_1 / \epsilon) \cos(2\pi x_2 / \epsilon) - 30 \right).
$$

Dividing the above equation for model problem 1 by $\epsilon$, we see that the problem does formally not fulfill the assumption $k^\epsilon = 1$ as assumed in Section 2. However, if we define $v^\epsilon(t, x) := u^\epsilon(\epsilon^{-1}t, x)$, then $v^\epsilon$ solves

$$
\epsilon \partial_t v^\epsilon - \epsilon \nabla \cdot (A'(x) \nabla v^\epsilon) + b'(x) \cdot \nabla v^\epsilon = 0 \quad \text{in } (0, 3\epsilon) \times \mathbb{R}^2 \quad \text{and} \quad v^\epsilon(0, \cdot) = v_0
$$

which fulfills the desired assumptions after dividing by $\epsilon$.

Model problem 1 deals with a stochastic perturbation of a space-periodic diffusion matrix. As a reference for the exact solution, we use the homogenized solution $u^0$, determined by ignoring the perturbation. In this example, we study in three test cases the properties of the HMM and of the global error indicator for increasing perturbations from the periodic setting. The first test is without perturbation,
Table 1. Model problem 1 with $\sigma = 0$. Relative error $\|e^N_u\|_{L^2(\mathbb{R}^2)}^{rel}$, relative estimated error $\Xi^N_{H,rel}$ and relative residual error indicator $\eta^N_{H,rel}$ for the last time step, $T = 3$, and various values of $(H, h)$. For further reference, each combination of $H, h$ is assigned to a fixed number in the column ‘uniform computation’. $i_{eff}$ denotes the efficiency index defined as the ratio between estimated error and real error.

| Uniform comp. | $H$  | $h$  | $\|e^N_u\|_{L^2(\mathbb{R}^2)}^{rel}$ | $\Xi^N_{H,rel}$ | $\eta^N_{H,rel}$ | $i_{eff}$ |
|---------------|------|------|-------------------------------------|-----------------|-----------------|----------|
| 1             | $2^{-3}$ | $2^{-7}$ | 0.037412                             | 0.293899        | 0.270058        | 7.86     |
| 2             | $2^{-3}$ | $2^{-4}$ | 0.018311                             | 0.148678        | 0.132966        | 8.12     |
| 3             | $2^{-4}$ | $2^{-7}$ | 0.012083                             | 0.091561        | 0.078663        | 7.58     |
| 4             | $2^{-5}$ | $2^{-8}$ | 0.005654                             | 0.047907        | 0.038703        | 8.47     |

Table 2. Model problem 1 with $\sigma = 0.01$. Errors and global error indicators and efficiency index for the last time step. Again we identify each value of $(H, h)$ with a fixed number for the corresponding uniform computation.

| Uniform comp. | $H$  | $h$  | $\|e^N_u\|_{L^2(\mathbb{R}^2)}^{rel}$ | $\Xi^N_{H,rel}$ | $\eta^N_{H,rel}$ | $i_{eff}$ |
|---------------|------|------|-------------------------------------|-----------------|-----------------|----------|
| 1             | $2^{-3}$ | $2^{-5}$ | 0.036343                             | 0.278529        | 0.261466        | 7.66     |
| 2             | $2^{-3}$ | $2^{-4}$ | 0.018509                             | 0.168850        | 0.15087         | 9.12     |
| 3             | $2^{-4}$ | $2^{-4}$ | 0.011993                             | 0.099409        | 0.081107        | 8.29     |
| 4             | $2^{-5}$ | $2^{-4}$ | 0.006588                             | 0.062704        | 0.044292        | 9.52     |

Table 3. Model problem 1 with $\sigma = 0.1$. Errors and global error indicators and efficiency index for the last time step.

| Uniform comp. | $H$  | $h$  | $\|e^N_u\|_{L^2(\mathbb{R}^2)}^{rel}$ | $\Xi^N_{H,rel}$ | $\eta^N_{H,rel}$ | $i_{eff}$ |
|---------------|------|------|-------------------------------------|-----------------|-----------------|----------|
| 1             | $2^{-3}$ | $2^{-5}$ | 0.036509                             | 0.290965        | 0.261406        | 7.97     |
| 2             | $2^{-3}$ | $2^{-4}$ | 0.018334                             | 0.182674        | 0.152931        | 9.96     |
| 3             | $2^{-4}$ | $2^{-4}$ | 0.011993                             | 0.099409        | 0.081107        | 8.29     |
| 4             | $2^{-5}$ | $2^{-4}$ | 0.006588                             | 0.062704        | 0.044292        | 9.52     |

i.e. $\sigma = 0$, the second with a small variance ($\sigma = 0.01$) and the third one with a relatively large variance ($\sigma = 0.1$). We expect the HMM to yield good results for all of the tests, but nevertheless the approximation should become worse with increasing variance. If the global error indicator is reliable in such examples, we expect it to show the same behaviour as the error itself, i.e. the experimental order of convergence (EOC) should be roughly the same for error and estimated error. At least, the estimated error must not converge faster than the error itself. For further tests in the periodic setting we refer to [37]. The number of time steps is set to $N = 20$ for all the computations. Hence, the time step size ($\Delta t = 0.15$) is quite large in comparison to the used mesh sizes $(H, h)$. This results in empirical convergence rates which do not reach second order. However, we have chosen this situation by purpose to see an influence of $\Delta t$, $H$ and...
Table 4. Model problem 1. EOC’s for errors and global error indicators. In the first and the second column we refer to the numbers of the computations that are used to determine the associated EOC’s. The numbers refer to the results depicted in Tab. 1, 2, and 3 (for $\sigma = 0, 0.01, 0.1$, respectively).

| Uniform comp. | $\sigma$ | EOC($\|e^N_{\sigma}(u)\|_{L^2(\mathbb{R}^2)}$) | EOC($\Xi^N_H$) | EOC($\eta^N_{H,rel}$) |
|---------------|---------|--------------------------------|----------------|----------------|
| 1$\rightarrow$3 | 0       | 1.6305                         | 1.6825         | 1.7795         |
| 2$\rightarrow$4 | 0       | 1.6954                         | 1.6339         | 1.7805         |
| 1$\rightarrow$3 | 0.01    | 1.5995                         | 1.4864         | 1.6887         |
| 2$\rightarrow$4 | 0.01    | 1.4903                         | 1.4291         | 1.7682         |
| 1$\rightarrow$3 | 0.1     | 1.5388                         | 1.3405         | 1.6394         |
| 2$\rightarrow$4 | 0.1     | 1.0444                         | 1.1774         | 1.6257         |

Table 5. Model problem 2. Error and estimated error for computations 0 to 4 on uniformly refined grids.

| Uniform comp. | $H$  | $h$  | $\|e^N_{\sigma}(u)\|_{L^2(\mathbb{R}^2)}$ | $\Xi_{H,rel}^{N,rel}$ |
|---------------|------|------|--------------------------------|----------------|
| 0             | 2$^{-5}$ | 2$^{-4}$ | 0.27411               | 0.91614         |
| 1             | 2$^{-2}$  | 2$^{-4}$ | 0.08612               | 0.64487         |
| 2             | 2$^{-4}$  | 2$^{-4}$ | 0.05191               | 0.32194         |
| 3             | 2$^{-2}$  | 2$^{-4}$ | 0.02726               | 0.18859         |
| 4             | 2$^{-5}$  | 2$^{-4}$ | 0.01245               | 0.09268         |

$h$ at the same time. The observations concerning the global error indicator and its speed of convergence are closer to practical applications and thus more relevant. Note that computations with smaller step sizes and $\sigma = 0$ yield the desired quadratic experimental orders of convergence, as demonstrated in [37].

Comparing the results depicted in Tables 1, 2 and 3 we see that the HMM solutions are very accurate even for the case of a quite large stochastic perturbation. This is also emphasized by Fig. 1 where a comparison between the isolines of the homogenized solution and the HMM approximation for $\sigma = 0.1$ is plotted. Nevertheless, we discover that the experimental orders of convergence (in space) are slightly decreasing for larger variance of the stochastic perturbation (cf. Tab. 4). Thus, the numerical experiments confirm our expectations. As a verification for the reliability of the a posteriori error estimate from Theorem 2.4, we observe that the estimated error captures the behaviour of the error itself. The estimated error is 8 to 9 times larger than the real error for all computations. These results indicate that the error estimator is of good quality, in the sense that it mimics the characteristics of the error itself. If the error worsens by the larger stochastic perturbation, the estimated error is doing the same. The reason for this behaviour is the approximation error indicator part of $\Xi^N_H$. The higher the stochastic perturbation, the longer it takes for this part to converge. This claim is also confirmed by Tab. 4. Here, the experimental orders of convergence for the estimated error without the approximation part $\eta^N_{H,rel}$ are constantly at about 1.7, independent of the variance $\sigma$. Hence, the approximation error part alone captures the influence of the perturbation and leads to reduced EOC’s of $\Xi^N_H$ with increasing variance.
Table 6. Model problem 2. Results of six numerical tests, using the adaptive strategy from Algorithm 1 for different tolerances and different $\sigma_{TOL}$. The adaptive results are related to the uniform results 2, 3 and 4 from Tab. 5.

| Adaptive comp. | $\|e^N_{a}||_{L^2((R^2))}$ | $\Xi^{n,ref}_H$ | Ref. uni. comp. | $\vartheta_e$ | $\vartheta_t$ |
|----------------|-----------------------------|-----------------|----------------|-------------|-------------|
| 1              | 0.04518                     | 0.21635         | 2              | 12.96%     | 19.38%     |
| 2              | 0.02724                     | 0.13866         | 3              | 0.07%      | 42.4%      |
| 3              | 0.0243                      | 0.13932         | 3              | 10.86%     | 38.59%     |
| 4              | 0.02298                     | 0.13679         | 3              | 15.7%      | 32.87%     |
| 5              | 0.0123                      | 0.08064         | 4              | 1.2%       | 25.58%     |
| 6              | 0.01081                     | 0.07942         | 4              | 13.17%     | 18.32%     |

Figure 2. Model problem 2. Adaptively refined grid and corresponding HMM approximation after the first time step for computation 6 of Tab. 6 (left), and at $t = 0.5$ for adaptive computation 1 (right).

Model problem 2. Find $u^* \in L^2(0, \frac{1}{2}; H^1(\mathbb{R}^2))$, with

$$\partial_t u^* - \varepsilon \nabla \cdot (A^*(x) \nabla u^*) + b^*(x) \cdot \nabla u^* = 0 \quad \text{in} \quad (0, \frac{1}{2}) \times \mathbb{R}^2 \quad \text{and}$$

$$u^*(0, x_1, x_2) = \begin{cases} 
4\cos(2\pi x_1) \sin(2^4 \pi x_2) & \text{in} \quad [-2^{-2}, 2^{-2}] \times [-9 \cdot 2^{-4}, -2^{-1}] \\
4\sin(2^2 \pi x_1) \sin(2^3 \pi x_2) & \text{in} \quad [-2^{-1}, -3 \cdot 2^{-3}] \times [3 \cdot 2^3, 4 \cdot 2^3] \\
4\sin(2^2 \pi x_1) \sin(2^3 \pi x_2) & \text{in} \quad [3 \cdot 2^{-3}, 2^{-1}] \times [3 \cdot 2^3, 4 \cdot 2^3] \\
0, & \text{else.}
\end{cases}$$

We choose $\varepsilon = 0.01$. The diffusion matrix $A^*(x)$ is given by

$$A^*(x) = \begin{pmatrix} 
\log \left( 4 + 2\sin^2(2\pi \frac{\log|1+x_1|}{\varepsilon}) \right) \\
0 \\
0 \\
\log \left( 4 + 2\cos^2(2\pi \frac{\log|1+x_2|}{\varepsilon}) \right)
\end{pmatrix}.$$ 

The advective term $b^*(x)$ is defined as

$$b^*(x) = \begin{pmatrix} 
\frac{3}{\varepsilon} \sin(2\pi \frac{x_1}{\varepsilon}) \sin(2\pi \frac{x_2}{\varepsilon}) + 2 \\
\frac{3}{\varepsilon} \cos(2\pi \frac{x_1}{\varepsilon}) \cos(2\pi \frac{x_2}{\varepsilon}) - 3
\end{pmatrix}.$$ 

Here, the diffusion matrix is rapidly oscillating, but not periodic. The exact solution $u^*$ is determined by a computation with a standard method on a highly resolved grid. We use $u^*(t, x + \frac{B(t)}{\varepsilon})$ as a reference for the accuracy of the HMM.
approximation $u_H$. In this example, we examine the adaptive Algorithm 1 that is based on the error indicators from Theorem 2.4. We performed six adaptive computations for various choices of $TOL$ and $\sigma_{TOL}$. For computations 1 and 2, we choose $\sigma_{TOL} = 1$ (equal distribution strategy), for computations 3, 4, 5, and 6 we only allow a variation of 80% of the average local error indicator, i.e. $\sigma_{TOL} = 0.8$. We start the algorithm with a uniform initial macro grid of mesh size $H = 2^{-2}$. For all computations, the micro mesh size is set to $h = 2^{-4}$ and the number of time steps is set to $N = 10$ which leads to $\Delta t = 0.05$. Tab. 5 shows that the average EOC of the HMM error is about 2.04, while the EOC for $\Xi_H$ tends to 1.79. The reason for not completely reaching second order can be found in the approximation error part and the part depending on $\Delta t$. Only in optimal cases and for very small time step sizes, we find the space EOC for the estimated error to be really quadratic.

Figures 2 and 3 show adaptively refined grids of computation 6 for $t = 0.05$ and $t = 0.5$, respectively. Moreover, Fig. 3 shows a perfect match of the isolines of the exact solution and a corresponding adaptive HMM approximation. Tab. 6 shows a comparison between computations on uniformly refined grids and comparable computations using the adaptive strategy. The results clearly indicate a significant advantage of the adaptive computations over the uniform one. We observe that the adaptive computations are at least 20% and up to 42% faster than comparable computations on uniform grids.

4. Proof of the a posteriori error estimate. In this section sketch the proof of the a posteriori error estimate in Theorem 2.4. The proof follows the ideas originally developed in [58, 36] for elliptic multiscale problems and is based on a reformulation of the numerical scheme in the case of periodic homogenization problems. A corresponding reformulation in the current setting has been derived in [37] and was exploited there to derive an a priori error estimate. The subsequent proof of the a posteriori error estimate will then be done in the two-scale variational setting and follows the classical concept for parabolic problems suggested e.g. in [60].

Let Assumption 2.2 be satisfied. Since the coefficients are then periodic in space, we use the notation $A_h(t, y) := A_h(t^n, yS)$ for $y \in S$. $b_h$ and $k_h$ are defined analogously. If $V_{\Delta t}$ denotes the piecewise constant functions in time, a discrete version
of the space $X$ is given by
\[
X_H(0,t^n) := V_{d,t}^0(0,t^n;V_H) \times L^2((0,t^n)\times \mathbb{R}^d,W_h(Y)) \text{ for } 1 \leq n \leq N
\]
and a discrete version of the two-scale operator $E$ is given by
\[
E_H(t)((u_H,u_h),(\Phi_H,\phi_h)) := E_H^{n+1}((u_H,u_h),(\Phi_H,\phi_h)) \text{ for } t \in (t^n,t^{n+1}],
\]
where
\[
E_H^n((u_H,u_h),(\Phi_H,\phi_h)) := \int_{\mathbb{R}^d} \int_Y (k_h(t^n,\cdot)\bar{b}_h(t^n) - b_h(t^n,\cdot)) \cdot \nabla_x \Phi_H u_h
- \int_{\mathbb{R}^d} \int_Y k_h(t^n,\cdot) \bar{b}_h(t^n) \cdot \nabla_x u_H \phi_h + \int_{\mathbb{R}^d} \int_Y b_h(t^n,\cdot) \cdot (\nabla_x u_H + \nabla_y u_H) \phi_h
+ \int_{\mathbb{R}^d} \int_Y A_h(t^n,\cdot) (\nabla_x u_H + \nabla_y u_H) \cdot (\nabla_x \Phi_H + \nabla_y \phi_h)
\]
for $(u_H,u_h),(\Phi_H,\phi_h) \in V_H \times L^2(\mathbb{R}^d,W_h(Y))$. In analogy to (4), we also define:
\[
G^N_H((u_H,u_h),(\Phi_H,\phi_h)) := \sum_{n=0}^{N-1} \left( \int_{t^n}^{t^{n+1}} (\partial_t u_H,\Phi_H)_{L^2(\mathbb{R}^d)} + E_H(t)((u_H,u_h),(\Phi_H,\phi_h)) \right)
+ \sum_{n=1}^{N-1} (u_H^n,\Phi^n_H)_{L^2(\mathbb{R}^d)} + ((u_H)_+,\Phi^0_H)_{L^2(\mathbb{R}^d)}.
\]
We suppose that the arguments belong to $X_H(0,T_0)$ and that $\partial_t (u_H^n)_{(t^n,t^{n+1})}$ exists. With these definitions, it has been shown in [37] that the following reformulation of the HMM as a discrete two-scale variational problem holds true.

**Theorem 4.1.** Assume $H \gg \epsilon$ and that $u_H^n, n = 0, \ldots, N,$ denote the HMM approximations given by Definition 2.1. Introducing $(u_H,K_h(u_H)) \in V_{d,t}^0(0,T_0;V_H) \times V_{d,t}^{0,0}(0,T_0) \times \mathbb{R}^d,W_h(Y))$ by

\[
u_{H,(t^n,t^{n+1})} := u_H^{n+1} \text{ and } K_h(u_H)_{(t^n,t^{n+1})} \times Y := \frac{1}{\epsilon} \left( R_H^{\epsilon}(u_H^n(\epsilon y) - u_H^n(\epsilon y)) \right),
\]
we then have that $(u_H,K_h(u_H))$ is a solution of
\[
G^N_H((u_H,K_h(u_H)),(\Phi_H,\phi_h)) = (v_H^0,(\Phi_H)_+^0)_{L^2(\mathbb{R}^d)}
\]
for all $(\Phi_H,\phi_h) \in X_H(0,T_0)$ and for all $N$, where $N \Delta t \leq T_0$.

In addition, under Assumption 2.2 it was proved in [37] that the HMM approximation $u_H$ converges to the homogenized solution $u_0$ strongly in $L^2$. This suggests comparing analytical and discrete two-scale equations

\[
G^N((u_0,u_1),(\Phi,\phi)) = (v_0,(\Phi)_+^0)_{L^2(\mathbb{R}^d)} \text{ and }
G^N_H((u_H,K_h(u_H)),(\Phi_H,\phi_h)) = (v_H^0,(\Phi_H)_+^0)_{L^2(\mathbb{R}^d)}
\]
and to use this variational setting for a-posteriori error estimation.

Analogous to the standard parabolic case (cf. [60]), we next define an appropriate dual backward problem that is then used for deriving an identity for the error $\|e^N\|^2_{L^2(\mathbb{R}^d)}$, $e^N := u^n(t^n,\cdot) - u_H^n$. 

**Definition 4.2.** We call \((z_0, z_1) \in X(0, t^N)\) the solution of the **dual backward problem**, if

\[
\int_0^{t^N} - (\Phi_0, \partial_t z_0)_{L^2(\mathbb{R}^d)} + E(t)((\Phi_0, \phi_1), (z_0, z_1)) = 0
\]

for all \((\Phi_0, \phi_1) \in L^2(0, t^N; H^1(\mathbb{R}^d)) \times L^2((0, t^N) \times \mathbb{R}^d, \tilde{H}_1^1(Y))\) and \(z_0(t^N, \cdot) = e^N\).

The dual backward problem 4.2 has a unique solution in \(X(0, t^N)\), since the homogenized problem (4) has a regular solution. Uniqueness and existence in the space \(L^2(0, t^N; H^1(\mathbb{R}^d)) \times L^2((0, t^N) \times \mathbb{R}^d, \tilde{H}_1^1(Y))\) can be achieved analogue to Theorem 3.1 in [38], since \((z_0, z_1)\) can be obtained as a two-scale homogenization limit of solutions to

\[
k^e \partial_t z^e + \nabla \cdot (A^e \nabla z^e) + \frac{1}{\epsilon} b^e \cdot \nabla z^e = 0 \text{ for } t < t^N \text{ and } z^e(t^N, \cdot) = e^N.
\]

**Remark 4.3.** As \((z_0, z_1) \in X(0, t^N)\) is the solution of the homogenized backward problem 4.2, \(z_0\) solves a macro problem of the form

\[
-A^e z_0 - \nabla \cdot (A^e \nabla z_0) = 0, \text{ for } t < t^N \text{ and } z_0(t^N, \cdot) = e^N.
\]

\(A^e\) is an elliptic, bounded and symmetric matrix in \((H^{1, \infty}(0, t^N))^{d \times d}\) and \(\|A^e\|_{H^{1, \infty}(0, t^N)}\) is bounded by a constant, only depending on \(A, b, k\).

We now derive an equation for the \(L^2\) error between \(u_0(t^N, \cdot)\) and \(u_H^N\).

**Lemma 4.4.** Let \((u_0, u_1) \in X(0, T_0)\) be the solution of problem (4), \((z_0, z_1) \in X(0, t^N)\) the solution of the dual backward problem 4.2, \(v_0\) the initial value of the problem and \(v_H\) a suitable FE-discretization of \(v_0\). Then the following error identity holds for every \((Z_H, z_h) \in Z_H(0, t^N)\):

\[
\|e^N\|_{L^2(\mathbb{R}^d)}^2 = G_H^N((u_H, K_h(u_H)), (Z_H - z_0, z_h - z_1)) + (v_0 - v_H, z_0(0, \cdot))_{L^2(\mathbb{R}^d)} + (G_H^N - G^N)((u_H, K_h(u_H)), (z_0, z_1)) + (v_H, z_0(0, \cdot) - (Z_H)_0^H)_{L^2(\mathbb{R}^d)}.
\]

**Proof.** For \(n \geq 0\) and \((Z_H, z_h) \in Z_H(0, t^N)\) we set \(Z_H^{n+1} := (Z_H)_0^H = (Z_H)^{n+1}_H\). We start with deriving

\[
G^N((u_0 - u_H, u_1 - K_h(u_H)), (z_0, z_1)) = (u_0(t^N, \cdot) - u_H^N, u_0(t^N, \cdot) - u_H^N)_{L^2(\mathbb{R}^d)}.
\]

Since \(u_0\) and \(z_0\) are continuous in time we have

\[
G^N((u_0, u_1), (z_0, z_1)) = \int_0^{t^N} ((\partial_t u_0, z_0)_{L^2(\mathbb{R}^d)} + E(t)((u_0, u_1), (z_0, z_1))) + (u_0(0, \cdot), z_0(0, \cdot))_{L^2(\mathbb{R}^d)}.
\]

Partial integration and \(z_0(t^N, \cdot) = u_0(t^N, \cdot) - u_H^N\) yields

\[
G^N((u_0, u_1), (z_0, z_1)) = (u_0(t^N, \cdot), u_0(t^N, \cdot) - u_H^N)_{L^2(\mathbb{R}^d)}.
\]
It remains to verify $G^N((u_H, K_h(u_H)), (z_0, z_1)) = ((u_H)^{N-1}, u_0(t^N, \cdot) - u_H^N)_{L^2(\mathbb{R}^d)}$. To do so we use the continuity of $z_0$:

$$G^N((u_H, K_h(u_H)), (z_0, z_1)) - \sum_{n=0}^{N-1} \left( \int_{t^n}^{t^{n+1}} (\partial_t u_H, z_0)_{L^2(\mathbb{R}^d)} + E(t)((u_H, K_h(u_H)), (z_0, z_1)) \right)$$

$$= \sum_{n=1}^{N-1} (u_H^{n+1} - u_H^n, z_0(t^n, \cdot))_{L^2(\mathbb{R}^d)} + (u_H^n, z_0(0, \cdot))_{L^2(\mathbb{R}^d)}$$

$$= - \sum_{n=0}^{N-1} (u_H^{n+1}, z_0(t^{n+1}, \cdot) - z_0(t^n, \cdot))_{L^2(\mathbb{R}^d)} + (u_H^n, z_0(t^n, \cdot))_{L^2(\mathbb{R}^d)}$$

$$= - \sum_{n=0}^{N-1} \left( \int_{t^n}^{t^{n+1}} (u_H, \partial_t z_0)_{L^2(\mathbb{R}^d)} \right) + (u_H^n, z_0(t^n, \cdot))_{L^2(\mathbb{R}^d)}.$$

Since $\partial_t u_H = 0$ on $(t^n, t^{n+1})$, we obtain from Definition 4.2

$$G^N((\Phi_H, \phi_h), (z_0, z_1)) = (u_H^n, z_0(t^n, \cdot))_{L^2(\mathbb{R}^d)}.$$

Together with (14) we get (13). Now, we are prepared for the error identity. With Theorem 2.3 and 4.1 we get for every $(Z_H, z_h) \in X_H(0, t^N)$:

$$G^N((u_H, K_h(u_H)), (Z_H, z_h)) - G^N((u_0, u_1), (Z_H, z_h)) = (v_H - v_0, Z_H^1)_{L^2(\mathbb{R}^d)} + (v_H - v_0, z_0(0, \cdot))_{L^2(\mathbb{R}^d)}.$$

(15)

Combining (13) and (15), we obtain:

$$\begin{align*}
(e^N, e_H^N)_{L^2(\mathbb{R}^d)} &= G^N((u_0, u_1), (z_0, z_1)) - G^N((u_H, K_h(u_H)), (z_0, z_1)) \\
&= G^N((u_0, u_1), (Z_H, z_h)) + G^N((u_H, K_h(u_H)), (Z_H, z_h)) \\
&+ G^N((u_H, K_h(u_H)), (z_0, z_1)) - G^N((u_H, K_h(u_H)), (z_0, z_1)) \\
&+ (v_H - v_0, z_0(0, \cdot) - Z_H^1)_{L^2(\mathbb{R}^d)} + (v_H - v_0, z_0(0, \cdot))_{L^2(\mathbb{R}^d)} \\
&= G^N((u_0, u_1), (z_0 - Z_H, z_1 - z_h)) \\
&+ G^N((u_H, K_h(u_H)), (Z_H - z_0, z_h - z_1)) \\
&+ (G^N_H - G^N)((u_H, K_h(u_H)), (z_0, z_1)) \\
&+ (v_H - v_0, z_0(0, \cdot) - Z_H^1)_{L^2(\mathbb{R}^d)} + (v_H - v_0, z_0(0, \cdot))_{L^2(\mathbb{R}^d)}.
\end{align*}$$

The equation $G^N((u_0, u_1), (z_0 - Z_H, z_1 - z_h)) = (v_0, z_0(0, \cdot) - Z_H^1)_{L^2(\mathbb{R}^d)}$ ends the proof.

In the following Lemma, we state some regularity estimates which are required to control the right hand side of (12).

**Lemma 4.5.** Under the general assumptions 2.2 we have that the solution $(z_0, z_1)$ of the dual backward problem 4.2 fulfills the estimates

$$| \int_{t^n}^{t^{n+1}} z_0(t, \cdot) dt |_{H^2(\mathbb{R}^d)} \leq C \int_{t^n}^{t^{n+1}} \| \partial_t z_0(t, \cdot) \|_{L^2(\mathbb{R}^d)} dt \quad \text{for } 0 \leq n < N - 1,$$

(16)
\[ | \int_{t^{n-1}}^{t^{n}} z_0(t, \cdot) \, dt |_{H^2(\mathbb{R}^d)} \leq C \Delta t \| e^N \|_{L^2(\mathbb{R}^d)} + C \| z_0(t^N, \cdot) - z_0(t^{N-1}, \cdot) \|_{L^2(\mathbb{R}^d)}, \]

(17)

\[ \int_0^{t^{N-1}} \| \partial_t z_0(t, \cdot) \|_{L^2(\mathbb{R}^d)} \leq C \left( \log \frac{t^N}{\Delta t} \right)^{\frac{1}{2}} \| e^N \|_{L^2(\mathbb{R}^d)}, \]

(18)

where \( C \) is a constant only depending on \( E \).

The proof of Lemma 4.5 adapts the ideas stated in the book of [60] for parabolic problems with time dependent coefficients and is left to the reader.

In the following we specify the test functions in equation (12) as

\[
Z_H(t, x)_{(|t^n, t^{n+1}|) = I_H(z_0^{n+1})(x), \text{ where } z_0^{n+1} := \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} z_0(t, \cdot) \, dt \text{ and } z_h(t, x, y) := I_h(z^1(t, x, \cdot))(y).}
\]

(19)

Here \( I_H \) and \( I_h \) denote corresponding Lagrange interpolation operators. This is not problematic since \((z_0, z_1)\) is sufficiently regular in \(1d, \ 2d \ \text{and} \ 3d.\) Therefore, \((Z_H, z_h)\) is an admissible test function in the error identity in Lemma 4.4.

Since Remark 4.3 shows that \( \| z_0(t, \cdot) \|_{L^2(\mathbb{R}^d)} \leq \| z_0(t^N, \cdot) \|_{L^2(\mathbb{R}^d)} \) for all \( t \leq t^N, \) we particularly have \( \| z_0(t^N, \cdot) - z_0(t^{N-1}, \cdot) \|_{L^2(\mathbb{R}^d)} \leq 2\| e^N \|_{L^2(\mathbb{R}^d)}. \) In the following we use this inequality without mentioning.

**Lemma 4.6.** Let \( I_H \) denote the Lagrange interpolation operator and let \( z_0^{n+1} \) be defined as in (19). Then we have the following estimate:

\[
\sum_{n=1}^{N-1} \left( \| u^n_H - u^{n-1}_H \|_{L^2(\mathbb{R}^d)} \right)^2 \leq C \left( \sum_{T \in T_H} \| u^n_H - u^{n-1}_H \|_{L^2(T)}^2 \right) \left( \log \frac{t^N}{\Delta t} \right)^{\frac{1}{2}} \| e^N \|_{L^2(\mathbb{R}^d)}. \]

(20)

We see that the expected second order convergence of the error estimator in (20) is slightly worsened by the factor \( \left( \log \frac{t^N}{\Delta t} \right)^{\frac{1}{2}} \) is \( O(\| \log(\Delta t) \|^{\frac{1}{2}}). \)

**Proof.** We estimate \( \| z_0^{n+1} \|_{H^2(\mathbb{R}^d)} \) by means of (16) and obtain for \( n \neq N - 1: \)

\[
\| z_0^{n+1} \|_{H^2(\mathbb{R}^d)} = \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} z_0(t, \cdot) \, dt \|_{H^2(\mathbb{R}^d)} \leq C \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \| \partial_t z_0(t, \cdot) \|_{L^2(\mathbb{R}^d)} \, dt
\]

and with (17)

\[
\| z_0^N \|_{H^2(\mathbb{R}^d)} \leq C \| e^N \|_{L^2(\mathbb{R}^d)} + \frac{1}{\Delta t} C \| z_0(t^N, \cdot) - z_0(t^{N-1}, \cdot) \|_{L^2(\mathbb{R}^d)}.
\]

Using these results and the approximation properties of the Lagrange interpolation operator yields the following:

\[
\sum_{n=1}^{N-1} \sum_{T \in T_H} \Delta t \left( \| u^n_H - u^{n-1}_H \|_{L^2(T)} \right) \leq C \left( \sum_{T \in T_H} \| u^n_H - u^{n-1}_H \|_{L^2(T)} \right) \left( \log \frac{t^N}{\Delta t} \right)^{\frac{1}{2}} \| e^N \|_{L^2(\mathbb{R}^d)}\]
Proof. Lemma 4.8. With (19), the following estimates hold true:

\[
\begin{align*}
&\leq C \sum_{n=1}^{N-2} \left( \sum_{T \in \mathcal{T}_H} \left\| \frac{u_H^{n+1} - u_H^n}{\Delta t} \right\|^2_{L^2(T)} \right)^{\frac{1}{2}} \int_{t^n}^{t^{n+1}} \left\| \partial_t z_0(t, \cdot) \right\|_{L^2(T)} \, dt \\
&+ C \left( \sum_{T \in \mathcal{T}_H} \left\| \frac{u_H^n - u_H^{n-1}}{\Delta t} \right\|^2_{L^2(T)} \right)^{\frac{1}{2}} \left( \left\| z_0(t^n, \cdot) - z_0(t^{n-1}, \cdot) \right\|_{L^2(\mathbb{R}^d)} + \Delta t \| e_N \|_{L^2(\mathbb{R}^d)} \right)
\end{align*}
\]

\[
\leq C \max_{1 \leq n \leq N-2} \left( \left( \sum_{T \in \mathcal{T}_H} \left\| \frac{u_H^{n+1} - u_H^n}{\Delta t} \right\|^2_{L^2(T)} \right)^{\frac{1}{2}} \left( \log \frac{t^n}{\Delta t} \right)^{\frac{1}{2}} \Delta t \| e_N \|_{L^2(\mathbb{R}^d)} \\
+ (2C + \Delta t) \left( \sum_{T \in \mathcal{T}_H} \left\| \frac{u_H^n - u_H^{n-1}}{\Delta t} \right\|^2_{L^2(T)} \right)^{\frac{1}{2}} \| e_N \|_{L^2(\mathbb{R}^d)}
\]

In the last step we made use of equation (18). For simplification, $2C + \Delta t$ is replaced by a sole constant $C$.

\[\square\]

Lemma 4.7. Let $z_0^{n+1}$ be given by (19), then the following estimate holds true:

\[
\sum_{n=0}^{N-1} \left( \| u_H \|_{n+1} - z_0(t^n, \cdot) \right)_{L^2(\mathbb{R}^d)}
\]

\[
\leq C \max_{1 \leq n \leq N-2} \left( \left( \sum_{T \in \mathcal{T}_H} \left\| \frac{u_H^{n+1} - u_H^n}{\Delta t} \right\|^2_{L^2(T)} \right)^{\frac{1}{2}} \left( \log \frac{t^n}{\Delta t} \right)^{\frac{1}{2}} \Delta t \| e_N \|_{L^2(\mathbb{R}^d)} \\
+ C \left( \sum_{T \in \mathcal{T}_H} \left\| \frac{u_H^n - u_H^{n-1}}{\Delta t} \right\|^2_{L^2(T)} \right)^{\frac{1}{2}} \Delta t \| e_N \|_{L^2(\mathbb{R}^d)}
\]

Proof. For $n \neq N - 1$ we calculate:

\[
\begin{align*}
&\left( u_H^{n+1} - u_H^n, z_0^n - z_0(t^n, \cdot) \right)_{L^2(\mathbb{R}^d)} \\
&\leq \Delta t \left\| u_H^{n+1} - u_H^n \right\|_{L^2(\mathbb{R}^d)} \left\| z_0^n - z_0(t^n, \cdot) \right\|_{L^2(\mathbb{R}^d)} \\
&\leq \Delta t \left\| u_H^{n+1} - u_H^n \right\|_{L^2(\mathbb{R}^d)} \int_{t^n}^{t^{n+1}} \left\| \partial_t z_0(t, \cdot) \right\|_{L^2(\mathbb{R}^d)} \, dt.
\end{align*}
\]

Proceeding analogously to the proof of Lemma 4.6 and using again equation (18) gives the result.

\[\square\]

Lemma 4.8. With (19), the following estimates hold true:

\[
\left( v_0 - v_H, z_0(0, \cdot) \right)_{L^2(\mathbb{R}^d)} \leq \left( \sum_{T \in \mathcal{T}_H} \left\| v_0 - v_H \right\|^2_{L^2(T)} \right)^{\frac{1}{2}} \| e_N \|_{L^2(\mathbb{R}^d)}, \quad (21)
\]
Proof. Proof of Theorem 2.4 With Lemma 4.4 we have the error identity

\[ \|e^N\|_{L^2(\mathbb{R}^d)} = \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E_H(t)((u_H, K_h(u_H)), (Z_H - z_0, z_h - z_1)) + (v_H, z_0(0, \cdot) - Z_H^1)_{L^2(\mathbb{R}^d)} + (v_0 - v_H, z_0(0, \cdot)_{L^2(\mathbb{R}^d)} \leq C \|e^N\|_{L^2(\mathbb{R}^d)} \]

is bounded by \( C \xi_H^N \|e^N\|_{L^2(\mathbb{R}^d)} \) For the rest we choose \((Z_H, z_h)\) such as in (19):

\[
\begin{align*}
\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E_H(t)((u_H, K_h(u_H)), (Z_H - z_0, z_h - z_1)) \\
+ \sum_{n=1}^{N-1} ([u_H]_n, (Z_H - z_0)^n)_{L^2(\mathbb{R}^d)} + (u_H^1, Z_H - z_0(0, \cdot))_{L^2(\mathbb{R}^d)} \\
+ (v_H, z_0(0, \cdot) - Z_H^1)_{L^2(\mathbb{R}^d)} + (v_0 - v_H, z_0(0, \cdot))_{L^2(\mathbb{R}^d)} \\
= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E_H(t)((u_H, K_h(u_H)), (Z_H - z_0, z_h - z_1)) \\
+ \sum_{n=1}^{N-1} ([u_H]_n, Z_H^{n+1} - z_0^{n+1})_{L^2(\mathbb{R}^d)} + \sum_{n=1}^{N-1} ([u_H]_n, z_0^{n+1} - z_0(t^n, \cdot))_{L^2(\mathbb{R}^d)} \\
+ (v_H - u_H^1, z_0(0, \cdot) - Z_H^1)_{L^2(\mathbb{R}^d)} + (v_0 - v_H, z_0(0, \cdot))_{L^2(\mathbb{R}^d)}.
\end{align*}
\]

The discrete time derivatives are estimated by using the Lemmata 4.6, 4.7 and 4.8. The term \( \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} E_H(t)((u_H, K_h(u_H)), (Z_H - z_0, z_h - z_1)) \) remains to be treated. For \( 0 < n \leq N \) we calculate
\[
\begin{align*}
&\cdot \nabla_y \left( \int_0^t \int_{t_n-1}^{t_n} z_h^n(x, x_T^n(y) - z_1(t, x, x_T^n(y)) \, dt \, dx \right) \, dy \\
= &\sum_{T \in T_H} \sum_{E \in \Gamma(T_h)} \int_E \left( \left[ A_h^n(t^n, x_T^n(y)) \nabla_x R_T^n(u_H^n) \right] \right)_{E_Y} \\
&\cdot \int_T \int_{t_n-1}^{t_n} z_h^n(x, x_T^n(y) - z_1(t, x, x_T^n(y)) \, dt \, dx \right) \, dy
\end{align*}
\]

With the approximation properties of the Lagrange operator we obtain:

\[
\sum_{n=1}^{N} \int_{t_n-1}^{t_n} \int_{\mathbb{R}^d} A_h \left( \nabla_x u_H^n + \nabla_y K_h^n(u_H^n) \right) \cdot \nabla_y (z_h - z_1) \, dy \, dx \quad (24)
\]

\[
\leq C \left( \sum_{n=1}^{N} \sum_{T \in T_H} \sum_{E \in \Gamma(T_h)} \triangle th_{E_Y} \| \left[ A_h^n(t^n, x_T^n(y)) \nabla_x R_T^n(u_H^n) \right] \right)_{E_Y} ^2 \| L^2(T \times E_Y) \right)^{\frac{1}{2}} \left| z_1 \right|_{L^2((0, t^N) \times \mathbb{R}^d; H^2(\mathbb{R}^d))}
\]

Due to periodicity of \( z_1(t, x, \cdot) \) we get \( \| z_1 \|_{L^2((0, t^N) \times \mathbb{R}^d; H^2(\mathbb{R}^d))} \leq C \| e_N \|_{L^2(\mathbb{R}^d)} \) by using (6). This ends the estimate for the part that depends on the term \( A_h (\nabla_x u_H^n + \nabla_y K_h^n(u_H^n)) \cdot \nabla_y (z_h - z_1) \). To derive an estimate for the \( A_h (\nabla_x u_H^n + \nabla_y K_h^n(u_H^n)) \cdot \nabla_x (Z_H - z_0) \)-part, we proceed as follows:

\[
\int_{t_n-1}^{t_n} \int_{\mathbb{R}^d} A_h(t^n, y) \left( \nabla_x u_H^n(x) + \nabla_y K_h^n(u_H^n)(x, y) \right) \cdot \nabla_x (Z_H - z_0)(t, x) \, dy \, dx \, dt
\]

\[
= \sum_{E \in \Gamma(T_h)} \int_E \int_Y \left[ A_h^n(t^n, x_T^n, y) \left( \nabla_x R_T^n(u_H^n) \right) \right]_{E_Y}(x) \left( \triangle t \right) Z_H^n - \int_{t_n-1}^{t_n} z_0(x) \, dy \, dx \, \text{d}x
\]

\[
\leq C \left( \sum_{E \in \Gamma(T_h)} H^3_E \| \int_Y \left[ A_h^n(t^n, \cdot) \nabla_x R_T^n(u_H^n) \right] \right)_{E_Y} ^2 \| L^2(E) \right)^{\frac{1}{2}} \left| \int_{t_n-1}^{t_n} z_0(t, \cdot) \, dt \right|_{H^2(\mathbb{R}^d)}.
\]

Here \( C \) depends on the maximal number of faces of an element of \( T_H \). Treating \( \left| \int_{t_n-1}^{t_n} z_0(t, \cdot) \, dt \right|_{H^2(\mathbb{R}^d)} \) as usual, we obtain

\[
\sum_{n=0}^{N-1} \int_{t_n-1}^{t_n+1} \int_{\mathbb{R}^d} A_h \left( \nabla_x u_H^n + \nabla_y K_h^n(u_H^n) \right) \cdot \nabla_x (Z_H - z_0)
\]

\[
\leq C \max_{1 \leq n \leq N-1} \left( \left( \sum_{E \in \Gamma(T_h)} H^3_E \| \int_Y \left[ A_h^n(\nabla_x u_H^n + \nabla_y K_h^n(u_H^n)) \right] \right)_{E_Y} ^2 \| L^2(E) \right)^{\frac{1}{2}} \left( \log \frac{t^N}{\Delta t} \right)^{\frac{1}{2}}
\]

\[
+ \left( \sum_{E \in \Gamma(T_h)} H^3_E \| \int_Y \left[ A_h^n(\nabla_x u_H^n + \nabla_y K_h^n(u_H^n)) \right] \right)_{E_Y} ^2 \| L^2(E) \right)^{\frac{1}{2}} \| e_N \|_{L^2(\mathbb{R}^d)}.
\]
Rewriting the right hand side with $x^*$ yields the estimate. And in complete analogy:

$$\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \int_Y (k_h \tilde{b}_h - b_h) \cdot \nabla_x (Z_H - z_0) K_h(u_H)$$

$$\leq C \max_{1 \leq n \leq N-1} \left( \left( \sum_{E \in \Gamma(T_H)} H_{E,H}^{3N,5,n,E,Y} \right)^{\frac{1}{2}} \left( \log \left( \frac{t_N}{\Delta t} \right) \right)^{\frac{1}{2}} \|e^N\|_{L^2(\mathbb{R}^d)} \right)$$

$$+ C \left( \sum_{E \in \Gamma(T_H)} H_{E,H}^{3N,5,n,E,Y} \right)^{\frac{1}{2}} \|e^N\|_{L^2(\mathbb{R}^d)}.$$

Similarly, using (6) and the approximation properties of the Lagrange operator:

$$\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \int_Y b_h(t^{n+1},y) \cdot \nabla_y K_h(u_H)(x,y) (z_h - z_1)(y) \ dy \ dx$$

$$\leq C \left( \sum_{n=1}^{N} \sum_{T \in T_H} \sum_{S \in T_h} \Delta t h_S^k \mu^{(3)}_{n,T,S} \right)^{\frac{1}{2}} \|e^N\|_{L^2(\mathbb{R}^d)}, \quad (25)$$

and

$$\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \int_Y (b_h(t^{n+1},y) - k_h(t^{n+1},y) \tilde{b}_h(t^{n+1})) \cdot \nabla_x u_H(x) (z_h - z_1)(y) \ dy \ dx$$

$$\leq C \left( \sum_{n=1}^{N} \sum_{T \in T_H} \sum_{S \in T_h} \Delta t h_S^k \mu^{(4)}_{n,T,S} \right)^{\frac{1}{2}} \|e^N\|_{L^2(\mathbb{R}^d)}, \quad (26)$$

Combining these results we get the desired a-posteriori error estimate for $\|u_0(t^N, \cdot) - u^N_H\|_{L^2(\mathbb{R}^d)}$ as claimed in Theorem 2.4. \qed

5. Conclusion. In this contribution we derived an a-posteriori error estimate for the heterogeneous multiscale FEM for advection-diffusion problems with rapidly oscillating coefficient functions and a large expected drift. This estimate was derived in order to establish a basis for error control and adaptive mesh refinement algorithms. In order to demonstrate the applicability and the advantages of the method, we performed two numerical experiments. In the first one, we analyzed the behaviour of the global error estimate in the case of an increasing stochastic perturbation from the periodic setting. In the second one we used an adaptive mesh refinement algorithm and analyzed its benefit over computations on uniformly refined grids. Even for the non-periodic diffusion matrix in model problem 2 good results could be shown.

REFERENCES

[1] A. Abdulle, Multiscale methods for advection-diffusion problems, Discrete and Continuous Dynamical Systems Series A, 5 (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, The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs, in Multiple scales problems in biomathematics, mechanics, physics and numerics, volume 31 of GAKUTO Internat. Ser. Math. Sci. Appl., pages 133–181. Gakkotosho, Tokyo, 2009.
[4] A. Abdulle and Y. Bai, Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems, *J. Comput. Phys.*, 231 (2012), 7014–7036.

[5] A. Abdulle and Y. Bai, Adaptive reduced basis finite element heterogeneous multiscale method, *Comput. Methods Appl. Mech. Engrg.*, 257 (2013), 203–220.

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

[7] A. Abdulle, W. E, B. Engquist and E. Vanden-Eijnden, The heterogeneous multiscale method, *Acta Numer.*, 21 (2012), 1–87.

[8] A. Abdulle and M. J. Grote, Finite element heterogeneous multiscale method for the wave equation, *Multiscale Model. Simul.*, 9 (2011), 766–792.

[9] A. Abdulle and M. E. Huber, Discontinuous Galerkin finite element heterogeneous multiscale method for advection-diffusion problems with multiple scales, *Numer. Math.*, 126 (2014), 589–633.

[10] A. Abdulle and A. Nonnenmacher, A posteriori error analysis of the heterogeneous multiscale method for homogenization problems, *C. R. Math. Acad. Sci. Paris*, 347 (2009), 1081–1086.

[11] A. Abdulle and A. Nonnenmacher, Adaptive finite element heterogeneous multiscale method for homogenization problems, *Comput. Methods Appl. Mech. Engrg.*, 200 (2011), 2710–2726.

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

[13] A. Abdulle and G. Vilmart, The effect of numerical integration in the finite element method for nonmonotone nonlinear elliptic problems with application to numerical homogenization methods, *C. R. Math. Acad. Sci. Paris*, 349 (2011), 1041–1046.

[14] 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).

[15] G. Allaire and A.-L. Raphael, Homogénéisation d’un modèle de convection-diffusion avec chimi/adsorption en milieu poreux, *Rapport Interne*, CMAP, Ecole Polytechnique, n. 604, 2006.

[16] 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), 521–528.

[17] 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).

[18] 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.

[19] R. Du and P. B. Ming, Convergence of the heterogeneous multiscale finite element method for elliptic problems with nonsmooth microstructures, *Multiscale Model. Simul.*, 8 (2010), 1770–1783.

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

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

[22] W. E and B. Engquist, The heterogeneous multi-scale method for homogenization problems, in *Multiscale methods in science and engineering*, volume 44 of *Lect. Notes Comput. Sci. Eng.*, pages 89–110. Springer, Berlin, 2005.

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

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

[25] Y. R. Efendiev, T. Y. Hou and X.-H. Wu, Convergence of a nonconforming multiscale finite element method, *SIAM J. Numer. Anal.*, 37 (2000), 888–910.

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

[27] B. Engquist, H. Holst and O. Runborg, Multiscale methods for wave propagation in heterogeneous media over long time, in *Numerical Analysis of Multiscale Computations*, Lecture Notes in Computational Science and Engineering, pages 167–186. Springer Verlag, 2012.

[28] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. I. A linear model problem, *SIAM J. Numer. Anal.*, 28 (1991), 43–77.

[29] K. Eriksson, C. Johnson and S. Larsson, Adaptive finite element methods for parabolic problems. VI. Analytic semigroups, *SIAM J. Numer. Anal.*, 35 (1998), 1315–1325 (electronic).
[30] F. R. F. Nataf and E. de Sturler, Optimal interface conditions for domain decomposition methods, CMAP (Ecole Polytechnique), (I.R. No. 301), 1994.

[31] A. Gloria, An analytical framework for the numerical homogenization of monotone elliptic operators and quasiconvex energies, Multiscale Model. Simul., 5 (2006), 996–1043 (electronic).

[32] A. Gloria, An analytical framework for numerical homogenization. II. Windowing and oversampling, Multiscale Model. Simul., 7 (2008), 274–293.

[33] A. Gloria, Reduction of the resonance error—Part 1: Approximation of homogenized coefficients, Math. Models Methods Appl. Sci., 21 (2011), 1601–1630.

[34] 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.

[35] L. Halpern, Artificial boundary conditions for the linear advection diffusion equation, Math. Comp., 46 (1986), 425–438.

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

[37] P. Henning and M. Ohlberger, The heterogeneous multiscale finite element method for advection-diffusion problems with rapidly oscillating coefficients and large expected drift, Netw. Heterog. Media, 5 (2010), 711–744.

[38] P. Henning and M. Ohlberger, A note on homogenization of advection-diffusion problems with large expected drift, Z. Anal. Anwend., 30 (2011), 319–339.

[39] P. Henning and M. Ohlberger, Error control and adaptivity for heterogeneous multiscale approximations of nonlinear monotone problems, Discrete Contin. Dyn. Syst. Ser. S, 8 (2015), 119–150.

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

[41] 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.

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

[43] T. Y. Hou, X.-H. Wu and Y. Zhang, Removing the cell resonance error in the multiscale finite element method via a Petrov-Galerkin formulation, Commun. Math. Sci., 2 (2004), 185–205.

[44] T. J. R. Hughes, Multiscale phenomena: Green’s functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, Comput. Methods Appl. Mech. Engrg., 127 (1995), 387–401.

[45] T. J. R. Hughes, G. R. Feijöo, L. Mazzei and J.-B. Quincy, The variational multiscale method - a paradigm for computational mechanics, Comput. Methods Appl. Mech. Engrg., 166 (1998), 3–24.

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

[47] V. V. Jikov, S. M. Kozlov and O. A. Oleinik, Homogenization of Differential Operators and Integral Functionals, Springer-Verlag, Berlin, 1994. Translated from the Russian by G. A. Yosifian.

[48] M. G. Larson and A. Målqvist, Adaptive variational multiscale methods based on a posteriori error estimation: duality techniques for elliptic problems, In Multiscale methods in science and engineering, volume 44 of Lect. Notes Comput. Sci. Eng., pages 181–193. Springer, Berlin, 2005.

[49] M. G. Larson and A. Målqvist, Adaptive variational multiscale methods based on a posteriori error estimation: energy norm estimates for elliptic problems, Comput. Methods Appl. Mech. Engrg., 196 (2007), 2313–2324.

[50] M. G. Larson and A. Målqvist, An adaptive variational multiscale method for convection-diffusion problems, Commun. Numer. Methods Engrg., 25 (2009), 65–79.

[51] R. Li, P. B. Ming and F. Tang, An efficient high order heterogeneous multiscale method for elliptic problems, Multiscale Model. Simul., 10 (2012), 259–283.

[52] A. Målqvist and D. Peterseim, Localization of elliptic multiscale problems, Math. Comp., 83 (2014), 2583–2603.

[53] 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.
[54] A.-M. Matache, Sparse two-scale FEM for homogenization problems, In Proceedings of the Fifth International Conference on Spectral and High Order Methods (ICOSAHOM-01) (Uppsala), 17 (2002), 659–669.

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

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

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

[58] 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).

[59] C. Schwab and A.-M. Matache, Generalized FEM for homogenization problems, In Multiscale and multiresolution methods, volume 20 of Lect. Notes Comput. Sci. Eng., pages 197–237. Springer, Berlin, 2002.

[60] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, volume 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1997.

Received January 2015; revised March 2015.

E-mail address: pathe@kth.se
E-mail address: mario.ohlberger@math.uni-muenster.de