Numerical homogenization method for parabolic advection-diffusion multiscale problems with large compressible flows

A. Abdulle,[1] M.E. Huber[1]

December 17, 2015

Abstract

We introduce a numerical homogenization method based on a discontinuous Galerkin finite element heterogeneous multiscale method (DG-HMM) to efficiently approximate the effective solution of parabolic advection-diffusion problems with rapidly varying coefficients, large Péclet number and compressible flows. To estimate the missing data of an effective model, numerical upscaling is performed which accurately captures the effects of microscopic solenoidal or gradient flow at a macroscopic scale such as enhancement or depletion of the effective diffusion. For compressible flow with periodic data, we derive sharp a priori error estimates for the macro and micro discretization errors which are robust in the advection dominated regime. Numerical tests confirm the error estimates for problems with periodic data and illustrate the applicability of our method for problems with non-periodic data.

Keywords: advection-diffusion equation, multiscale method, numerical homogenization, discontinuous Galerkin methods, compressible flows, high Péclet number.

AMS subject classification (2010): 65N12, 35K15, 74Q10, 76N99

1 Introduction

Transport processes driven by advection and diffusion are used to model a variety of phenomena in natural sciences, e.g., transport of passive tracers in the atmosphere (meteorology), propagation of pollutants in groundwater (hydrology) or separation of bio molecules (biology), see [5, 44, 49]. In many of those applied problems one is interested to predict the behavior of the transported entity over long time and at a scale of interest in space that can be several orders of magnitude larger than the small length scale ε at which the governing processes take place. This typically leads to parabolic advection-diffusion problems with highly heterogeneous data varying at scale ε and a large Péclet number of size $\mathcal{O}(\varepsilon^{-1})$. Thus there are two major issues when using standard numerical methods like the finite element method (FEM) for spatial discretization. First, to adequately capture the correct effective behavior the small scale ε has to be resolved by the spatial mesh (scale resolution) which yields prohibitively high computational cost. Second, the large Péclet number causes poor stability of the method.

Homogenization theory is the usual tool to rigorously derive effective equations for partial differential equations (PDEs) driven by microscopic heterogeneities, e.g., see [18, 39]. For parabolic advection-diffusion PDEs with stationary data, Péclet number of size $\mathcal{O}(\varepsilon^{-1})$ and *general* advection, i.e., neither zero mean nor divergence free flow, most of the results available in literature have been derived for periodic data and for the infinite spatial domain $\mathbb{R}^d$. In [38, 43] (see [39, Chapter 2] for a summary), homogenization results with explicit formulas for the effective diffusion and drift are given. It is shown that in coordinates moving with the macroscopic drift the effective behavior as $\varepsilon \to 0$ is described by the solution of a purely diffusive homogenized parabolic PDE. Those results have been generalized in [45] to problems with nonlinear advection, in [15] to porous media problems and in [26] to problems with time dependent data (note that reactive terms are allowed in the results of [15, 26, 43]). Further, in [45] the concept of two-scale convergence introduced in [46] has been adapted to the situation of large advection to justify multiscale expansions in Lagrangian coordinates. Beyond the periodic setting in the spatial domain $\mathbb{R}^d$, difficulties occur due to macroscopic variations of advection and the interplay of the large

---

[1]ANMC, Mathematics Section, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland; Assyr.Abdulle@epfl.ch, Martin.Huber@epfl.ch.

drift with boundary conditions. In [13] problems with locally periodic data and non-symmetric diffusion tensors are studied and in [14], for periodic data, the asymptotic profile (with respect to ε) of the finescale solutions on a bounded domain with homogeneous Dirichlet boundary conditions is described (classical homogenization would yield $u^\varepsilon \to 0$). We note that the effects of advection on micro scale onto macro dynamics are manifold already in the periodic setting. Advection at the micro scale can either enhance (for solenoidal flows) or deplete (for gradient flows) macro diffusion and can generate a macroscopic drift which in general differs from the arithmetic mean, e.g., see [44, 51].

To numerically approximate effective solutions of PDEs with data varying at multiple length scales numerical homogenization methods have been developed. For parabolic advection-diffusion PDEs with Péclet number of order $\mathcal{O}(\varepsilon^{-1})$ but *divergence free* advection, multiscale methods have been proposed in [1, 33, 47]. In [1] a method to solve advection-diffusion problems with constant scalar diffusion is introduced, where the advection is obtained as the flux of an elliptic diffusion multiscale equation which is solved using a finite element heterogeneous multiscale method (FE-HMM), see [28] and [2, 3] for general reviews of heterogeneous multiscale methods (HMM) and FE-HMM, respectively, and stabilized explicit Runge-Kutta methods (ROCK) [8] serve as time integrator. Note that the analysis given in [1] focuses on the FE-HMM used for the flux computation, i.e., an elliptic multiscale problem. To solve advection dominated problems numerically, discontinuous Galerkin finite element methods (DG-FEM) have become increasingly popular due to their stability and conservation properties. Further the block diagonal mass matrix of DG-FEM makes them convenient for time dependent problems. Initially developed in [34] for first order hyperbolic problems, DG methods have been extended to elliptic problems (see [16]) and to advection-diffusion problems, see [36] and [30, 31] for elliptic and parabolic problems, respectively. Recently, in [7], a discontinuous Galerkin heterogeneous multiscale method (DG-HMM) has been introduced and analyzed for elliptic advection-diffusion multiscale problems with general velocity field and possibly large Péclet number however bounded independently of ε.

We describe now the two main numerical strategies and convergence results of numerical homogenization methods for the full time dependent advection-diffusion equation. In [33], the implicit Euler scheme in time is combined with an FE-HMM in space. The method is formulated in Lagrangian coordinates moving in the direction of an upscaled drift, i.e., their effective equation is purely diffusive. The effective data is estimated by solving elliptic advection-diffusion problems within microscopic domains. However, in the Lagrangian coordinates, the microscopic simulations depend on time even if the original data is time independent. Thus they have to be reevaluated at each time step, which is computationally expensive. For periodic data and divergence free velocity field, fully discrete space-time a priori error estimates in Lagrangian coordinates are presented, i.e., there are no advective terms at macro scale in their error analysis. Further, a posteriori error control is discussed in [32]. In [48] a space-discrete multiscale method extending the multiscale finite element method (MsFEM) from [12] has been proposed (see [35] for an MsFEM applied to cellular flow). The usual finite element basis functions are replaced by solutions to elliptic advection-diffusion problems on patches of macroscopic size with Dirichlet boundary conditions. As the Péclet number within the localized problems is still large, the basis functions in general have boundary layers. For periodic data, an a priori error analysis with explicit rates for the convergence of the numerical solution towards the finescale solution is given in [47]. However, the convergence rates are not robust for $\varepsilon \to 0$. As for [33], divergence free velocity fields are assumed in the analysis.

In this paper, we propose a discontinuous Galerkin (DG) HMM for the parabolic advection-diffusion equation with data rapidly varying at scale ε. Following the concept of HMM, we use DG-FEM to discretize an effective parabolic advection-diffusion equation in space whose a priori unknown data (diffusion tensor and advection field) are estimated by numerical upscaling. Microscopic simulations using standard FEM are performed within sampling domains of size $\mathcal{O}(\varepsilon)$ by solving elliptic advection-diffusion problems which are indefinite (as compressible flows are allowed) and whose differential operators have non-trivial kernels spanned by a positive function, i.e., the analytical framework is the Fredholm alternative. The effective data then is approximated by suitable averaging of the microscopic FEM solutions. The proposed multiscale strategy is well-defined for general data without assuming any particular structure of their spatial heterogeneities (like periodicity or random stationarity) and leads to computational cost independent of the small scale ε as the sampling domains are of microscopic size. We thus avoid scale resolution by resorting to numerical upscaling and get a stable numerical scheme as the DG-FEM at macro scale is well-suited for advection dominated problems.

We first summarize the main novelties of our method and analysis. The proposed numerical upscaling procedure captures the effective dynamics for *general* flows b^ε as neither $\operatorname{div} b^\varepsilon = 0$ nor zero mean for b^ε is assumed, while the methods proposed in [33, 47] are only designed for incompressible flows b^ε. Our main a priori error estimates are also valid for periodic flows b^ε without assuming $\operatorname{div} b^\varepsilon = 0$ and they are

robust with respect to the Péclet number which scales as $\mathcal{O}(\varepsilon^{-1})$. In contrast, the estimates in [33, 47] are only valid for *incompressible* periodic flows. Further, in [33] a priori error estimates are derived in a moving frame (i.e., an analysis of a pure diffusion problem) while the a priori error bounds in [47] are not robust for $\varepsilon \to 0$ (as only the right-hand side of their estimates has terms scaling as $\mathcal{O}(\varepsilon^{-1})$). The cost of the numerical upscaling procedure is independent of the smallest scale and for time independent velocity and permeability tensors, the micro problems used to recover the effective coefficients have to be done only once. In contrast, the method [47] has a total complexity of $\mathcal{O}(\varepsilon^{-d})$ in an offline stage to compute effective basis functions (a computation that can however be localized within each macro element). As our method is formulated in natural physical variables, the micro simulations are independent of time while in [33] these computations must be performed at each time step, unless the multiscale coefficients are independent of the macro variable. Finally we mention that for the methods proposed in [33, 47] as for our method, the micro computations are independent one from another and can be performed in parallel.

We briefly also discuss the main difficulties in analyzing our new method. We first mention that the main convergence result of this article is a convergence analysis for problems on rectangular spatial domains with periodically oscillating data and periodic boundary conditions, which is the setting for which rigorous homogenization is available and thus studying the convergence of the numerical solution towards an effective solution makes sense. To obtain robust convergence rates of the spatial L^2 norm in the advection dominated regime, the use of a "weighting function" which monotonically decreases along the flow has proved to be an essential tool, see [17, 40] (and [7] for a multiscale context). Such a tool is however not available on periodic domains. As a remedy, we periodize the weighting function such that the analysis of [17, 40] remains valid up to a small boundary layer. A peculiarity of our results is, that the L^2 error at the boundary of the macro domain cannot be controlled robustly with respect to ε in the advection dominated regime. We emphasize that the same technical difficulty occurs when extending the singlescale analysis of DG-FEM given in [17, 40] for Dirichlet and Neumann boundary conditions to advection-diffusion problems with *periodic* boundary conditions. This issue is thus not directly related to our multiscale approach. Our a priori error estimates are valid for periodic flows b^ε without assuming $\operatorname{div} b^\varepsilon = 0$. Allowing $\operatorname{div} b^\varepsilon \neq 0$ yields non-symmetric indefinite differential operators in space and thus theoretical results have to be derived in the setting of the Fredholm alternative, while for $\operatorname{div} b^\varepsilon = 0$ (like in [33, 47]) the differential operators in space are coercive and one can work in the standard setting of the Lax-Milgram theorem. In our error analysis, variational crimes due to numerical upscaling have to be controlled. Compared to our previous work [7] about stationary problems with Péclet number independent of ε^{-1} we need a more refined analysis of those non-consistent perturbations as we now consider problems with Péclet number of size $\mathcal{O}(\varepsilon^{-1})$. Further, while in [7] pure diffusion problems were used for micro sampling, the multiscale method proposed in this article is based on advection-diffusion problems at micro scale and thus new estimates for the micro error are derived. For the sake of simplicity, we solve the indefinite micro problems by the standard FEM. Note that the Péclet number for the elliptic advection-diffusion problems within the micro domains is of size $\mathcal{O}(1)$ with respect to ε, as the sampling domains are of size ε. Hence, the micro problems are in general diffusion dominated and there are no stability issues for standard FEM. Nonetheless, specialized numerical schemes tailored for this class of non-coercive problems could easily be used instead. We mention for instance monotone and stabilized FEM, finite volume methods (FVM) or combined FEM-FVM, see [20, 23, 42, 55].

The outline of the article is as follows. In Section 2 we introduce the model problem and an associated effective equation and summarize the homogenization results for periodic data. The space-discrete multiscale method is then introduced in Section 3 and existence and uniqueness of the numerical solution are shown. In Section 4, the main results about the convergence of the spatial macro and micro error for periodic data are formulated and they are proved subsequently in Section 5. Extensive numerical tests to validate the theoretical convergence rates and to illustrate the applicability of the multiscale method beyond the periodic setting are provided in Section 6. Finally, a conclusion is given in Section 7.

Notations. We denote the usual Sobolev spaces by $W^{k,p}(\Omega)$ and write $H^k(\Omega)$ for $p = 2$. For bounded Ω, the subspace $H_0^1(\Omega)$ of $H^1(\Omega)$ contains the functions with zero trace on $\partial\Omega$ and, for $Y = (0,1)^d$, $H_{per}^k(Y)$ is defined as the closure in $H^k(Y)$ of $\mathcal{C}_{per}^\infty(Y)$ (smooth periodic functions on Y). If Ω is a cuboid, $H_{per}^k(\Omega)$ is analogously defined. For $T > 0$ and a Banach space X, the spaces $\mathcal{C}^0([0,T], X)$ and $L^2(0, T; X)$ are given by the set of functions $g \colon [0,T] \to X$ that are continuous or L^2 integrable, respectively. We use $e_1, \ldots, e_d$ for the canonical basis of $\mathbb{R}^d$ and $|b|$ denotes the Euclidean norm for $b \in \mathbb{R}^d$. For $a \in \mathbb{R}^{d \times d}$ (with identity matrix denoted by Id) we write $a_{ij} = a e_j \cdot e_i$ and $\|a\|_{\mathcal{F}}$ for its Frobenius norm.

2 Model problem and its homogenization

We consider the parabolic multiscale advection-diffusion problem

$$\partial_t u^\varepsilon(x,t) - \mathrm{div}(a^\varepsilon(x)\nabla u^\varepsilon(x,t)) + b^\varepsilon(x) \cdot \nabla u^\varepsilon(x,t) = f(x), \qquad \text{in } \Omega \times (0,T),$$
$$u^\varepsilon(x,0) = g(x), \qquad \text{in } \Omega, \tag{1}$$

with either $\Omega = \mathbb{R}^d$ or a convex polygonal domain $\Omega \subset \mathbb{R}^d$ (for $d \in \{2,3\}$), $T > 0$, diffusion tensor $a^\varepsilon(x) \in (L^\infty(\Omega))^{d \times d}$, velocity field $b^\varepsilon(x) \in (L^\infty(\Omega))^d$, source term $f \in L^2(\Omega)$, initial condition $g \in L^2(\Omega)$ and suitable boundary conditions if $\Omega \neq \mathbb{R}^d$. The small parameter $\varepsilon > 0$ denotes a microscopic scale at which the data a^ε and b^ε vary rapidly. We assume that the tensors a^ε are uniformly elliptic and bounded and allow vector fields b^ε of magnitude $\mathcal{O}(\varepsilon^{-1})$, i.e., for any $\varepsilon > 0$ and almost every (a.e.) $x \in \Omega$ we have

$$a^\varepsilon(x)\xi \cdot \xi \geq \lambda |\xi|^2, \quad |a^\varepsilon(x)\xi| \leq \Lambda |\xi|, \quad |b^\varepsilon(x)| \leq \mathcal{B}\varepsilon^{-1}, \qquad \forall \xi \in \mathbb{R}^d, \tag{2}$$

where $0 < \lambda \leq \Lambda$ and $\mathcal{B} > 0$. Note that we *neither* assume that b^ε is divergence free *nor* that it has zero macroscopic mean. Given a Hilbert space $H \subseteq H^1(\Omega)$, incorporating eventual boundary conditions of (1) and with dual space denoted by H^*, existence and uniqueness of a weak solution u^ε to (1) is obtained in

$$E = \{v \in L^2(0,T;H) \,|\, \partial_t v \in L^2(0,T;H^*)\}, \tag{3}$$

e.g., see [57, Corollary 23.26], as due to (2) the spatial differential operators satisfy the Gårding inequality

$$\int_\Omega a^\varepsilon(x)\nabla v \cdot \nabla v \, dx + \int_\Omega b^\varepsilon(x) \cdot \nabla v \, v \, dx \geq \frac{\lambda}{2}\|v\|_{H^1(\Omega)}^2 - \frac{\lambda}{2}\left(1 + \frac{\mathcal{B}^2}{\lambda^2 \varepsilon^2}\right)\|v\|_{L^2(\Omega)}^2, \tag{4}$$

for $v \in H^1(\Omega)$. Standard a priori bounds for u^ε in the $\mathcal{C}^0([0,T], L^2(\Omega))$ and $L^2(0,T;H^1(\Omega))$ norm can be derived using (4) but the bounds generally depend on negative powers of ε, i.e., are not uniform in ε.

Remark 2.1. The model problem (1) with its velocity field of order $\mathcal{O}(\varepsilon^{-1})$ is obtained by introducing the macroscopic variables $x = \varepsilon y$ and $t = \varepsilon^2 \tau$, called *diffusive* or *parabolic* scaling, for an unscaled parabolic advection-diffusion problem at microscopic scale (with variables denoted by (y, τ)). Note that the alternative scaling $x = \varepsilon y$, $t = \varepsilon \tau$, called advective or hyperbolic scaling, would yield a purely hyperbolic effective equation neglecting any effects of diffusion, e.g., see [51, Chapter 13].

Effective model. In this article we assume the existence of an upscaled model (associated to (1))

$$\partial_t u^{\mathrm{eff}}(x,t) - \mathrm{div}(a^{\mathrm{eff}}(x)\nabla u^{\mathrm{eff}}(x,t)) + b^{\mathrm{eff}}(x) \cdot \nabla u^{\mathrm{eff}}(x,t) = f(x), \qquad \text{in } \Omega \times (0,T),$$
$$u^{\mathrm{eff}}(x,0) = g(x), \qquad \text{in } \Omega, \tag{5}$$

with effective data $a^{\mathrm{eff}} \in (L^\infty(\Omega))^{d \times d}$, $b^{\mathrm{eff}} \in (L^\infty(\Omega))^d$ and effective solution u^{eff} where the rapid oscillations (of period ε) are averaged out. To ensure existence and uniqueness of u^{eff} we assume that a^{eff} and b^{eff} satisfy (2) with possibly altered constants, i.e., b^{eff} can still be of order $\mathcal{O}(\varepsilon^{-1})$. If a^ε, b^ε are periodic, the existence of such effective model (5) can be proved by homogenization.

Homogenization for periodic data. We summarize the homogenization results derived in [15, 26, 38, 43, 45] that are relevant in our context.

Let $\Omega = \mathbb{R}^d$ and $f \equiv 0$ in (1). Assume that a^ε and b^ε are periodic, i.e., $a^\varepsilon(x) = a(x/\varepsilon)$ and $b^\varepsilon(x) = \varepsilon^{-1}b(x/\varepsilon)$ where $a(y)$ and $b(y)$ are Y-periodic and sufficiently smooth. The homogenization of (1) is formally obtained by using the multiscale expansion in moving coordinates

$$u^\varepsilon(x,t) = u^0(x - b^{\mathrm{eff}}t, \tfrac{x}{\varepsilon}, t) + \varepsilon u^1(x - b^{\mathrm{eff}}t, \tfrac{x}{\varepsilon}, t) + \dots, \qquad \text{with } b^{\mathrm{eff}} = \varepsilon^{-1}b^*, \tag{6}$$

where $u^k(x, \cdot, t)$ are Y-periodic and $b^* \in \mathbb{R}^d$ is an a priori unknown drift. Plugging (6) into equation (1) and sorting the terms with respect to their powers in ε yields a cascade of equations $\mathcal{L}u^k = f^{(k)}$ on Y with $k \in \mathbb{N}$, operator $\mathcal{L} = -\mathrm{div}(a(y)\nabla_y \cdot) + b(y) \cdot \nabla_y$, maps $f^{(k)} \colon \Omega \times Y \times (0,T) \to \mathbb{R}$ and $f^{(0)} \equiv 0$.

As $\mathcal{L}$ is non-symmetric and indefinite, the framework to study solvability of $\mathcal{L}u^k = f^{(k)}$ is the Fredholm alternative. First, for $k = 0$, we get that $u^0(x,y,t) = u^0(x,t)$ as due to the maximum principle we have $\ker \mathcal{L} = \mathbb{R}$. Then, the Fredholm alternative further ensures that $\dim \ker \mathcal{L}^* = 1$ for the adjoint operator $\mathcal{L}^*$. Hence, there is a unique (normalized) solution to the adjoint problem: find $\rho \in H^1_{per}(Y)$ with $\int_Y \rho(y)dy = 1$ such that

$$\int_Y a(y)\nabla z \cdot \nabla \rho \, dy + \int_Y b(y) \cdot \nabla z \, \rho \, dy = 0, \qquad \forall z \in H^1_{per}(Y). \tag{7}$$

Due to the maximum principle ρ is strictly positive and, if $\operatorname{div} b = 0$, we have $\rho \equiv 1$.

Second, for $k = 1$, the equation $\mathcal{L}u^1 = f^{(1)}$ is separated into the cell problems: for $1 \le i \le d$, find $\chi^i \in H^1_{per}(Y)$ with $\int_Y \chi^i dy = 0$ such that

$$\int_Y a(y)(e_i + \nabla\chi^i) \cdot \nabla z \, dy + \int_Y b(y) \cdot (e_i + \nabla\chi^i)z \, dy = \int_Y b^* \cdot e_i z \, dy, \quad \forall z \in H^1_{per}(Y). \tag{8}$$

According to the compatibility condition of the Fredholm alternative a solution χ^i exists if and only if $b^* \cdot e_i = \int_Y [b(y) \cdot e_i - \operatorname{div}_y(a(y)e_i)]\rho(y)dy$. Further, the normalization $\int_Y \chi^i dy = 0$ yields uniqueness of the solution. Similarly, the compatibility condition for $k = 2$ holds if and only if u^0 solves the problem $\partial_t u^0 - \operatorname{div}(a^{\text{eff}}\nabla u^0) = 0$ with a^{eff} given in (9) below. In particular, the drift b^* introduces the suitable set of Lagrangian coordinates such that u^0 can be described by a pure diffusion process.

Finally, setting $u^{\text{eff}}(x,t) = u^0(x - b^{\text{eff}}t, t)$ one finds, that u^{eff} solves the effective problem (5) with tensor $a^{\text{eff}} \in \mathbb{R}^{d \times d}$ and drift $b^{\text{eff}} \in \mathbb{R}^d$ explicitly given by

$$a^{\text{eff}}_{ij} = \int_Y a(y)(e_j + \nabla\chi^j) \cdot e_i \, \rho \, dy + \int_Y \operatorname{div}[a(y)e_i\chi^j] \, \rho \, dy - \int_Y (b(y) - b^*) \cdot e_i \chi^j \rho \, dy, \tag{9a}$$

$$b^{\text{eff}} = \varepsilon^{-1}b^*, \qquad b^* = \int_Y [b(y) - (\operatorname{div}_y(a(y)e_1), \ldots, \operatorname{div}_y(a(y)e_d))^T]\rho \, dy, \tag{9b}$$

where ρ and χ^k (for $1 \le k \le d$) solve the cell problems (7) and (8), respectively. Further, the finescale solutions u^ε converge strongly in $L^2(0, T; L^2(\mathbb{R}^d))$ towards the effective solution u^{eff} and the values of a^{eff} and b^{eff} are independent of the normalization of χ^k and ρ.

Note, that even if the data only has the regularity $a(y) \in (L^\infty(Y))^{d \times d}$ and $b(y) \in (L^\infty(Y))^d$, existence and uniqueness of the solutions to the cell problems (8) and (7) are still ensured and thus the formulas (9) for the homogenized data are well-defined, see [26, Proposition 1]. Sufficient regularity is only needed to prove ellipticity of a^{eff} and the convergence of u^ε towards u^0.

Finally, we discuss the effective diffusion tensor a^{eff} defined in (9a). While the first term in (9a) is known from diffusion problems (but now weighted by ρ), the second and the third terms account for the effects of micro advection onto macro diffusion and are referred to as depletion and enhancement of the effective diffusion, respectively, e.g., see [51, Section 13.6]. Ellipticity of a^{eff} can be shown by using that $a^{\text{eff}}\xi \cdot \xi = \hat{a}^{\text{eff}}\xi \cdot \xi$ (for all $\xi \in \mathbb{R}^d$) with $\hat{a}^{\text{eff}}$ given by

$$\hat{a}^{\text{eff}}_{ij} = \int_Y a(y)(e_j + \nabla\chi^j) \cdot (e_i + \nabla\chi^i) \, \rho \, dy, \tag{10}$$

whose ellipticity follows from (2) and the property $\rho > 0$. Further, as $\int_{\mathbb{R}^d}(a^{\text{eff}} - \hat{a}^{\text{eff}})\nabla v \cdot \nabla w \, dx$ for all $v, w \in H^1(\mathbb{R}^d)$, replacing a^{eff} by $\hat{a}^{\text{eff}}$ in (5) yields an equivalent formulation of the effective PDE which is commonly used in periodic homogenization, e.g., see [15, 21, 26].

Alternative model problem. Instead of (1) one could consider the advection-diffusion process

$$\partial_t \tilde{u}^\varepsilon(x,t) - \operatorname{div}(a^\varepsilon(x)\nabla\tilde{u}^\varepsilon(x,t) + b^\varepsilon(x)\tilde{u}^\varepsilon(x,t)) = f, \qquad \text{in } \Omega \times (0, T), \tag{11}$$

– a formulation motivated by the conservation of mass. For periodic data $a^\varepsilon, b^\varepsilon$, $\Omega = \mathbb{R}^d$ and $f \equiv 0$, homogenization results have as well been derived in [15, 26, 38, 43].

While problems (1) and (11) share many features, there is one major difference. In contrast to (1), where the solutions u^ε can be approximated by a non-oscillating function u^{eff}, the solutions $\tilde{u}^\varepsilon$ to (11) have oscillations of order $\mathcal{O}(1)$, which therefore cannot be neglected. In particular, one can show that

$$\tilde{u}^\varepsilon(x,t) \approx \tilde{u}^{\text{eff}}(x,t)\rho(x/\varepsilon),$$

where ρ solves (7) and $\tilde{u}^{\text{eff}}$ is the solution to an effective advection-diffusion problem. We emphasize that the multiscale method developed in this article could easily be adapted to approximate $\tilde{u}^{\text{eff}}$.

3 Multiscale method

We now consider the model problem (1) on a bounded, open cuboid $\Omega \subset \mathbb{R}^d$ with periodic boundary conditions on the boundary $\partial\Omega$ and prescribe periodic initial conditions $g \in L^2_{per}(\Omega)$. We emphasize that the multiscale method can straightforwardly be formulated for general bounded domains Ω and different boundary conditions (like Dirichlet, Neumann or mixed boundary conditions).

Before defining the numerical homogenization method in Section 3.2, we first give in Section 3.1 the necessary tools for the macro and micro solvers. Existence and uniqueness of the numerical solution of the new multiscale method is then proved in Section 3.3.

3.1 Spatial macro and micro discretization

In this section, we introduce the discretization of the macro and the micro spatial domains.

Macro discretization. Consider a shape-regular partition $\mathcal{T}_H$ of the cuboid Ω where $\mathcal{T}_H$ consists of open simplicial elements $K \in \mathcal{T}_H$ with straight edges. The index H denotes the mesh size $H = \max_{K \in \mathcal{T}_H} H_K$ where $H_K = \operatorname{diam} K$. The partition $\mathcal{T}_H$ is chosen independently of ε, i.e., $H \gg \varepsilon$ is allowed and $\mathcal{T}_H$ is thus called macro mesh. We denote by $\mathcal{E}$ the set of all $(d-1)$-dimensional interfaces of the elements $K \in \mathcal{T}_H$. The diameter of an edge $e \in \mathcal{E}$ is further denoted by H_e and we use the notation $\int_\Gamma \cdot = \sum_{e \in \mathcal{E}} \int_e \cdot$. To keep the exposition simple, we suppose that $\mathcal{T}_H$ is conformal, i.e., it has no hanging nodes. Further, we assume that element interfaces at opposite boundaries of $\partial\Omega$ coincide and thus can be identified as one single interface. Therefore, every $e \in \mathcal{E}$ belongs to exactly *two* distinct elements $K_1, K_2 \in \mathcal{T}_H$.

Let $v \in H^1(\mathcal{T}_H)$ be in the piecewise Sobolev space

$$H^1(\mathcal{T}_H) = \prod_{K \in \mathcal{T}_H} H^1(K) = \{v \in L^2(\Omega) \,|\, v \in H^1(K) \text{ for all } K \in \mathcal{T}_H\}, \tag{12}$$

and $e \in \mathcal{E}$ be shared by the macro elements $K_1, K_2 \in \mathcal{T}_H$. For $j = 1, 2$, we denote by n_j and v_j the outer normal vector and the trace of v on e, respectively, associated to the element K_j. We define the normal jump and the average of v on the edge e by

$$[\![v]\!] = v_1 n_1 + v_2 n_2, \qquad \{v\} = \frac{1}{2}(v_1 + v_2),$$

which are defined analogously for vector-valued functions $v \in (H^1(\mathcal{T}_H))^d$. For $b \in \mathbb{R}^d$ we set

$$\begin{cases} K_i = K_1, \\ K_o = K_2, \end{cases} \text{if } b \cdot n_1 < 0, \qquad \text{or} \qquad \begin{cases} K_i = K_2, \\ K_o = K_1, \end{cases} \text{if } b \cdot n_1 > 0,$$

and call K_i and K_o the inflow and outflow element with respect to b, respectively. For convenience, we still use the notation K_i, K_o even if $b \cdot n_1 = 0$. We then choose K_i as the element among K_1 and K_2 whose barycenter is "smaller" with respect to the lexicographical order on $\mathbb{R}^d$. We further use the notation $v_i = v_{K_i}$, $n_i = n_{K_i}$ and $v_o = v_{K_o}$, $n_o = n_{K_o}$.

On the macro mesh $\mathcal{T}_H$ we introduce the *discontinuous* finite element space

$$V^1(\Omega, \mathcal{T}_H) = \{v^H \in L^2(\Omega) \,|\, v^H \in \mathcal{P}^1(K), \forall K \in \mathcal{T}_H\}, \tag{13}$$

where $\mathcal{P}^1(K)$ is the space of affine polynomials on $K \in \mathcal{T}_H$.

To evaluate integrals on macro elements $K \in \mathcal{T}_H$ we use the barycentric quadrature formula

$$\int_K p(x)dx \approx |K| p(x_K), \qquad \text{for } p \in \mathcal{C}^0(\overline{K}), \tag{14}$$

where x_K denotes the barycenter of the element $K \in \mathcal{T}_H$.

Remark 3.1. The generalization of the multiscale method to non-conformal macro meshes, i.e., meshes with hanging nodes, is straightforward. At the same time, the condition that interfaces on opposite sides of $\partial\Omega$ have to match can be dropped.

Micro discretization. Our multiscale method is based on numerically upscaled data obtained by local micro sampling. For each macro element $K \in \mathcal{T}_H$ we define the micro sampling domain $K_\delta = x_K + \delta(-1/2, 1/2)^d$ of size $\delta \geq \varepsilon$ around the barycenter x_K. We discretize the sampling domain K_δ by a shape-regular, conformal, simplicial partition $\mathcal{T}_h$ whose elements $T \in \mathcal{T}_h$ have straight edges and diameter denoted by h_T. The micro mesh size h is then given by $h = \max_{T \in \mathcal{T}_h} h_T$. On $\mathcal{T}_h$, we consider the space of *continuous* finite element functions

$$S^1(K_\delta, \mathcal{T}_h) = \{v^h \in W(K_\delta) \,|\, v^h \in \mathcal{P}^1(T), \forall T \in \mathcal{T}_H\}, \tag{15}$$

where $\mathcal{P}^1(T)$ is the set of affine polynomials on $T \in \mathcal{T}_h$ and $W(K_\delta) \subseteq H^1(K_\delta)$ is a closed subspace.

3.2 Space-discrete DG-HMM

In this section, we propose a space-discrete numerical homogenization method to solve (1) by coupling a discontinuous Galerkin finite element method at macro scale with finite element simulations within micro sampling domains.

Multiscale method. Let $u_0^H \in V^1(\Omega, \mathcal{T}_H)$ be given. The multiscale method is then defined as: find $u^H \colon [0,T] \to V^1(\Omega, \mathcal{T}_H)$ with $u^H(x,0) = u_0^H$ such that

$$\int_\Omega \partial_t u^H(t)\, w^H dx + B(u^H(t), w^H) = \int_\Omega f\, w^H dx, \quad \forall\, w^H \in V^1(\Omega, \mathcal{T}_H), t \in (0,T), \qquad (16)$$

where the macro bilinear form $B(v^H, w^H)$ for $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$ is given by the sum $B = B_D + B_A$ of the diffusion and advection forms (recall that $\int_\Gamma \cdot = \sum_{e \in \mathcal{E}} \int_e \cdot$)

$$B_D(v^H, w^H) = \sum_{K \in \mathcal{T}_H} |K| a_K^{\mathrm{eff},h} \nabla v^H(x_K) \cdot \nabla w^H(x_K) - \int_\Gamma \{a_K^{\mathrm{eff},h} \nabla v^H\} \cdot [\![w^H]\!] ds + \int_\Gamma \mu [\![v^H]\!] \cdot [\![w^H]\!] ds,$$

$$B_A(v^H, w^H) = \sum_{K \in \mathcal{T}_H} |K| b_K^{\mathrm{eff},h} \cdot \nabla v^H(x_K)\, w^H(x_K) - \int_\Gamma \{b_K^{\mathrm{eff},h}\} \cdot [\![v^H]\!] w_i^H\, ds,$$

$$(17)$$

where $a_K^{\mathrm{eff},h}$ and $b_K^{\mathrm{eff},h}$ given in (21) and (19), respectively, are obtained by micro sampling. For $e \in \mathcal{E}$, the penalty function μ is chosen as $\mu|_e = \alpha H_e^{-1} \|\{a_K^{\mathrm{eff},h}\}\|_\mathcal{F}$ with a penalization parameter $\alpha > 1$ and the inflow trace w_i^H is taken with respect to $\{b_K^{\mathrm{eff},h}\}$.

Micro solver. Let $K \in \mathcal{T}_H$ and K_δ be its associated sampling domain. The micro sampling strategy consists of two steps.

Step 1. Find $\rho_K^h \in S^1(K_\delta, \mathcal{T}_h)$ with $\int_{K_\delta} \rho_K^h dx = |K_\delta|$ such that

$$\int_{K_\delta} a^\varepsilon \nabla z^h \cdot \nabla \rho_K^h dx + \int_{K_\delta} b^\varepsilon \cdot \nabla z^h \rho_K^h dx = 0, \qquad \forall\, z^h \in S^1(K_\delta, \mathcal{T}_h), \qquad (18)$$

and compute the effective drift $b_K^{\mathrm{eff},h}$ given by

$$b_K^{\mathrm{eff},h} = \frac{1}{|K_\delta|} \int_{K_\delta} (a^\varepsilon)^T \nabla \rho_K^h dx + \frac{1}{|K_\delta|} \int_{K_\delta} b^\varepsilon\, \rho_K^h dx. \qquad (19)$$

Step 2. For $1 \le i \le d$, find $\psi_K^{i,h} \in S^1(K_\delta, \mathcal{T}_h)$ with $\int_{K_\delta} \psi_K^{i,h} dx = 0$ such that

$$\int_{K_\delta} a^\varepsilon (e_i + \nabla \psi_K^{i,h}) \cdot \nabla z^h dx + \int_{K_\delta} b^\varepsilon \cdot (e_i + \nabla \psi_K^{i,h}) z^h dx = \int_{K_\delta} b_K^{\mathrm{eff},h} \cdot e_i z^h dx, \qquad \forall\, z^h \in S^1(K_\delta, \mathcal{T}_h),$$

$$(20)$$

with $b_K^{\mathrm{eff},h}$ the effective drift (19) from *Step 1*. The effective diffusion tensor $a_K^{\mathrm{eff},h}$ is then defined as

$$a_K^{\mathrm{eff},h} = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon (I + \nabla \psi_K^h)\, \rho_K^h dx - \frac{1}{|K_\delta|} \int_{K_\delta} (a^\varepsilon)^T \nabla \rho_K^h\, \psi_K^h dx + \frac{1}{|K_\delta|} \int_{K_\delta} (b_K^{\mathrm{eff},h} - b^\varepsilon)\, \psi_K^h \rho_K^h dx, \ (21)$$

where $\psi_K^h(x) = (\psi_K^{1,h}(x), \dots, \psi_K^{d,h}(x)) \in \mathbb{R}^{1 \times d}$, $\nabla \psi_K^h(x) = (\nabla \psi_K^{1,h}(x), \dots, \nabla \psi_K^{d,h}(x)) \in \mathbb{R}^{d \times d}$ and ρ_K^h is the micro function solving (18) computed in *Step 1*.

Coupling conditions. To determine the coupling between the macro and the micro solver boundary conditions for the micro problems (18) and (20) are encoded into $W(K_\delta) \subseteq H^1(K_\delta)$. In this article, we consider the periodic coupling $W(K_\delta) = H^1_{per}(K_\delta)$. If the data a^ε and b^ε are periodic we further choose a sampling domain size $\delta \ge \varepsilon$ satisfying $\delta/\varepsilon \in \mathbb{N}_{>0}$.

Our multiscale method is however not restricted to periodic coupling. Recall that for pure diffusion problems Dirichlet and Neumann coupling conditions are commonly used, e.g., see [56]. Note that Dirichlet coupling $W(K_\delta) = H^1_0(K_\delta)$ is not possible for the considered model problem, as then the adjoint micro problem (18) has only the trivial solution $\rho_K^h \equiv 0$, which violates the normalization constraint $\int_{K_\delta} \rho_K^h dx = |K_\delta|$. Further, Neumann coupling does not immediately generalize to the micro problems (18) and (20) as they cannot be written as minimization problems and as the compatibility condition of the Fredholm alternative has to be respected, see [37, Section 4.6.2] for a detailed discussion.

Discussion of the multiscale method. In practice, only one stiffness matrix has to be assembled to solve the micro problems (18) and (20) as the problem (18) is the adjoint problem (with zero source term) of the problem (20). Further, the d micro problems (20) indexed by $1 \le i \le d$ only differ in their right-hand side term. We emphasize that the micro sampling can easily be parallelized as the micro problems in different sampling domains K_δ are completely independent.

7

Further, the effective drift defined in (19) can be represented alternatively by

$$b_K^{\mathrm{eff},h} \cdot e_i = \frac{1}{|K_\delta|} \int_{K_\delta} b^\varepsilon \cdot (e_i + \nabla \psi_K^{i,h}) dx, \qquad 1 \le i \le d, \tag{22}$$

which can be obtained by taking $z^h \equiv 1$ in (20). Note that the representation (22) cannot be used in an implementation as the value of $b_K^{\mathrm{eff},h}$ is needed to solve the equation (20) in order to compute $\psi_K^{i,h}$. However, we recover the same formula for the effective drift as used for the stationary advection-diffusion problems with *small* drift $b^\varepsilon = \mathcal{O}(1)$ studied in [7].

Remark 3.2. As the quadrature formula (14) is exact for affine polynomials and we only consider the DG space $V^1(\Omega, \mathcal{T}_H)$ of piecewise affine functions, an *equivalent* formulation of the bilinear forms B_D and B_A defined in (17) is obtained by writing their first terms (the ones with a quadrature formula) as integral

$$\sum_{K \in \mathcal{T}_H} \int_K a_K^{\mathrm{eff},h} \nabla v^H(x) \cdot \nabla w^H(x) dx, \qquad \sum_{K \in \mathcal{T}_H} \int_K b_K^{\mathrm{eff},h} \cdot \nabla v^H(x) w^H(x) dx.$$

We will use the integral formulation in Section 5 as it is more suitable for the analysis.

Remark 3.3. For periodic a^ε and b^ε, i.e., if there exist $a(y) \in (L^\infty(Y))^{d \times d}$ and $b(y) \in (L^\infty(Y))^d$ both Y-periodic in y such that $a^\varepsilon(x) = a(x/\varepsilon)$ and $b^\varepsilon(x) = \varepsilon^{-1} b(x/\varepsilon)$, respectively, the micro problems (18) and (20) are independent of the macro element $K \in \mathcal{T}_H$. In this case, we omit the index K used to denote their solutions $\psi_K^{i,h} = \psi^{i,h}$, $\rho_K^h = \rho^h$ and the effective data $a_K^{\mathrm{eff},h} = a^{\mathrm{eff},h}$, $b_K^{\mathrm{eff},h} = b^{\mathrm{eff},h}$.

Further, motivated by (10), one might consider another upscaled diffusion tensor (instead of $a^{\mathrm{eff},h}$)

$$\hat{a}^{\mathrm{eff},h} = \frac{1}{|K_\delta|} \int_{K_\delta} (Id + \nabla \psi^h)^T a^\varepsilon(x)(Id + \nabla \psi^h) \rho^h dx, \tag{23}$$

where $\psi^h = (\nabla \psi^{1,h}, \dots, \nabla \psi^{d,h}) \in \mathbb{R}^{d \times d}$ with $\psi^{i,h}$ and ρ^h solving (20) and (18), respectively. Note that in contrast to the continuous setting, cf. (10), in general $a^{\mathrm{eff},h} \xi \cdot \xi \ne \hat{a}^{\mathrm{eff},h} \xi \cdot \xi$ for $\xi \in \mathbb{R}^d$, but the equality is only recovered in the limit as $h \to 0$.

3.3 Existence and uniqueness of the numerical solution

In this section, we prove that the multiscale method (16) is well-defined if the micro meshes $\mathcal{T}_h$ are sufficiently fine and adapted to the diffusion tensor $a^\varepsilon(x)$.

Macro equation. As the multiscale method (16) is discrete in space, but continuous in time (a method of lines) the macro equation (16) is a linear ordinary differential equation (ODE) for the unknown $u^H : [0, T] \to V^1(\Omega, \mathcal{T}_H)$. Hence, assuming that

(M) the micro problems (18) and (20) are uniquely solvable for all $K \in \mathcal{T}_H$, $1 \le i \le d$,

which can be proved in some situations (see Lemmas 3.4 and 3.5), the macro map B given in (17) is well-defined and global existence and uniqueness of u^H follows from the Picard-Lindelöf theorem.

Micro equations. To study the existence and uniqueness of a solution to the micro problems (18) and (20), i.e., condition **(M)**, we rephrase them as systems of linear equations, see (24), and apply Fredholm's theorem in linear algebra, see (25).

Let $K \in \mathcal{T}_H$ and K_δ be its associated sampling domain. Assume that the periodic boundary conditions for the micro problems (18) and (20) are strongly enforced by using micro meshes $\mathcal{T}_h$ such that the degrees of freedom on opposite faces of K_δ are identified as one single degree of freedom. Let $M = \dim S^1(K_\delta, \mathcal{T}_h)$, $\beta_1^h, \dots, \beta_M^h$ be the nodal basis functions and write $\beta^h(x) = (\beta_1^h(x), \dots, \beta_M^h(x))^T$. We define the micro stiffness matrix $S \in \mathbb{R}^{M \times M}$ and the right-hand side vectors $r^{(i)} = (r_1^{(i)}, \dots, r_d^{(i)})^T \in \mathbb{R}^M$, for $1 \le i \le d$,

$$S_{kl} = \int_{K_\delta} a^\varepsilon \nabla \beta_l^h \cdot \nabla \beta_k^h dx + \int_{K_\delta} b^\varepsilon \cdot \nabla \beta_l^h \, \beta_k^h dx, \tag{24a}$$

$$r_k^{(i)} = \int_{K_\delta} (b_K^{\mathrm{eff},h} - b^\varepsilon) \cdot e_i \beta_k^h dx - \int_{K_\delta} a^\varepsilon e_i \cdot \nabla \beta_k^h dx, \tag{24b}$$

with $1 \le k, l \le M$. Using the coefficient vectors $\boldsymbol{\rho}_K, \boldsymbol{\psi}_K^i \in \mathbb{R}^M$ via the identification $\rho_K^h(x) = \boldsymbol{\rho}_K \cdot \beta^h(x)$, $\psi_K^{i,h}(x) = \boldsymbol{\psi}_K^i \cdot \beta^h(x)$, the micro problems (18) and (20) can be written as: find $\boldsymbol{\rho}_K, \boldsymbol{\psi}_K^i \in \mathbb{R}^M$ such that

$$S^T \boldsymbol{\rho}_K = 0, \text{ with } \int_{K_\delta} \rho_K^h dx = |K_\delta|, \qquad S \boldsymbol{\psi}_K^i = r^{(i)}, \text{ with } \int_{K_\delta} \psi_K^{i,h} dx = 0. \tag{25}$$

To prove unique solvability of (25), we follow [22] where the standard finite element method for indefinite advection-diffusion problems with the Laplace operator $-\Delta\cdot$ as diffusion operator and Neumann boundary conditions is studied. Compared to [22], only minor modifications are necessary to adopt their proof to problems with general diffusion operator $-\operatorname{div}(a\nabla\cdot)$ and periodic boundary conditions.

Lemma 3.4. *Let the set* $\operatorname{adj}(\mathcal{T}_h)$ *be defined as*

$$\operatorname{adj}(\mathcal{T}_h) = \{(k,l) \in \{1,\dots,M\} \times \{1,\dots,M\} \,|\, k \neq l \text{ and } \operatorname{supp}(\beta_k^h \beta_l^h) \neq \emptyset\}, \tag{26}$$

which can be identified as the set of pairs of adjacent vertices in $\mathcal{T}_h$ *when interpreting the mesh* $\mathcal{T}_h$ *as undirected graph. Assume that* $\mathcal{T}_h$ *is quasi-uniform, e.g., see [24, Condition (3.2.28)], and that there exists* $\nu > 0$ *independent of* h, δ *and* ε *such that*

$$\int_{K_\delta} a^\varepsilon(x) \nabla \beta_l^h \cdot \nabla \beta_k^h dx \leq -\nu h^{d-2}, \qquad \forall\, (k,l) \in \operatorname{adj}(\mathcal{T}_h). \tag{27}$$

Then, there exists $\hat{h}_0 > 0$ *independent of* ε *such that for all* h *with* $h/\varepsilon < \hat{h}_0$ *condition* (**M**) *holds, i.e., the micro problems* (18) *and* (20) *have a unique solution* ρ_K^h *and* $\psi_K^{i,h}$, *respectively.*

Proof. The linear equations in (25) are uniquely solvable – by Fredholm's theorem in linear algebra – if

$$\operatorname{rank} S = M - 1, \qquad r^{(i)} \in (\ker S^T)^\perp, \tag{28}$$

as the uniqueness of the solution then follows from the normalization constraints imposed in (25). The proof of (28) follows the proof of [22, Proposition 3], see [37, Lemma 4.2.4] for the details. $\qquad\square$

Let us briefly discuss the assumptions of Lemma 3.4. First, the result is robust in ε as the micro mesh size $h \leq \delta$ scales as $\mathcal{O}(\varepsilon)$ and thus $h/\varepsilon = \mathcal{O}(1)$. Second, assumption (27) postulates that the finite element space $S^1(K_\delta, \mathcal{T}_h)$ fits the structure of the tensor a^ε. Conditions similar to (27) (mostly for $a^\varepsilon = Id$) have been used in the context of discrete maximum principles for finite element methods, e.g., see [25] for an early work and [41] for references. For piecewise linear simplicial finite elements ($\mathcal{P}^1$-FEM) applied to PDEs with diffusion terms $-\operatorname{div}(a(x)\nabla\cdot)$, where $a(x) > 0$ is scalar, explicit geometric conditions on the mesh $\mathcal{T}_h$ exist to ensure (27). In particular, it is sufficient that the interior angles between the $(d-1)$-dimensional interfaces of the simplices $T \in \mathcal{T}_h$ are non-obtuse, i.e., smaller or equal to $\pi/2$ (for elliptic diffusion PDEs) or strictly acute, i.e., smaller than $\pi/2 - \epsilon$ for some $\epsilon > 0$ (for elliptic diffusion-reaction PDEs). We are however not aware of any result for general matrix-valued diffusion *tensors* $a(x) \in \mathbb{R}^{d \times d}$.

In Lemma 3.5, we summarize the result which is relevant for advection-diffusion PDEs.

Lemma 3.5. *Let the tensor* $a^\varepsilon(x)$ *be given by* $a^\varepsilon(x) = \gamma^\varepsilon(x) Id$, *where* $\gamma^\varepsilon \colon \overline{\Omega} \to \mathbb{R}$ *with* $\gamma^\varepsilon(x) \geq \lambda > 0$ *and* $\gamma \in L^\infty(\Omega)$. *If the micro mesh* $\mathcal{T}_h$ *is shape-regular, quasi-uniform and all angles between its* $(d-1)$-*dimensional interfaces are smaller than* $\pi/2 - \epsilon$ *for some* $\epsilon > 0$, *i.e., strictly acute, then the hypothesis* (27) *holds with a constant* ν *only depending on* λ, ϵ, *the shape-regularity and the quasi-uniformity of* $\mathcal{T}_h$.

Proof. The proof is based on the arguments given in, e.g., [41, Section 5.1]. We refer to [37, Lemma 4.2.5] for a detailed proof. $\qquad\square$

4 Main results

In this section, we present the main results about convergence of the numerical approximation u^H obtained by the multiscale method (16) towards an effective solution u^{eff}.

We provide a rigorous convergence analysis for the setting where homogenization results for the model problem (1) are available (periodic a^ε and b^ε and vanishing source f) and consider the multiscale method (16) with optimal coupling. We refer to this setting as *periodic setting* described by the hypotheses

$$a^\varepsilon(x) = a\Big(\frac{x}{\varepsilon}\Big), \quad b^\varepsilon(x) = \frac{1}{\varepsilon} b\Big(\frac{x}{\varepsilon}\Big), \quad f \equiv 0, \quad W(K_\delta) = H_{per}^1(K_\delta), \quad \delta = \varepsilon, \tag{29}$$

where $a(y) \in (L^\infty(Y))^{d \times d}$ and $b(y) \in (L^\infty(Y))^d$ are Y-periodic in y. Like in Section 3, we consider the model problem (1) and its effective equation (5) on a cuboid Ω with periodic boundary conditions on $\partial\Omega$ and initial condition $g \in L_{per}^2(\Omega)$. In what follows, we compare the numerical solution obtained by (16) to u^{eff} which solves the effective problem (5) with data defined in (9).

Due to the periodicity of the data the micro solutions $\psi_K^{i,h}$, ρ_K^h to (18) and (20), respectively, and the upscaled data $a_K^{\text{eff},h}$ and $b_K^{\text{eff},h}$ defined in (21) and (19), respectively, are independent of $K \in \mathcal{T}_H$. In Sections 4 and 5 we therefore omit the subscript K and use the simplified notation $\rho^h = \rho_K^h$, $\psi^{i,h} = \psi_K^{i,h}$, $a^{\text{eff},h} = a_K^{\text{eff},h}$ and $b^{\text{eff},h} = b_K^{\text{eff},h}$, see Remark 3.3.

Interior of Ω. For $\tau > 0$ we define the subdomain $\Omega_\tau \subset \Omega$ and its complement Ω_τ^c by

$$\Omega_\tau = \{x \in \Omega \mid \text{dist}(x, \partial\Omega) > \tau\}, \qquad \Omega_\tau^c = \Omega \setminus \overline{\Omega}_\tau, \tag{30}$$

where $\text{dist}(x, \partial\Omega) = \inf_{y \in \partial\Omega} |x - y|$. We assume for simplicity that the macro mesh $\mathcal{T}_H$ is aligned to the subdomain Ω_τ, i.e., for $K \in \mathcal{T}_H$ we have either $K \cap \Omega_\tau = \emptyset$ or $K \cap \Omega_\tau^c = \emptyset$.

DG norm. For $v \in H^1(\mathcal{T}_H)$ defined in (12), we introduce the diffusive and advective DG norms $\|v\|_D$ and $\|v\|_{A,\tau}$ given by

$$\|v\|_D^2 = \|\nabla v\|_{L^2(\Omega)}^2 + |v|_{*,D}^2, \qquad \|v\|_{A,\tau}^2 = \|v\|_{L^2(\Omega_\tau)}^2 + |v|_{*,A}^2, \tag{31}$$

with the jump norms $|\cdot|_{*,\cdot}$.

$$|v|_{*,R}^2 = \sum_{e \in \mathcal{E}} \left\| \gamma(R)^{1/2} [\![v]\!] \right\|_{L^2(e)}^2, \qquad \text{where } \gamma(R) = \begin{cases} \alpha H_e^{-1}, & R = D, \\ |\varepsilon b^{\text{eff},h} \cdot n|, & R = A, \end{cases} \tag{32}$$

with $b^{\text{eff},h}$ defined in (19). Note that $\gamma(A)$ is scaled such that $\gamma(A) = \mathcal{O}(1)$ with respect to ε (see Lemma 5.3 for a proof).

The advective norm $\|\cdot\|_{A,\tau}$ does *not* control the L^2 norm on the whole domain Ω but only on the interior subdomain Ω_τ defined in (30), see Remark 5.1 for discussion. Further, compared to our previous work [7], the advective jump quantified by $|\cdot|_{*,A}$ is measured with respect to the numerical drift $b^{\text{eff},h}$ and *not* the exact drift b^{eff}. This definition of $|\cdot|_{*,A}$ is in the spirit of [19] where DG for first order hyperbolic problems with stabilization different from traditional upwinding has been studied.

HMM error. To quantify the error due to the HMM upscaling procedure we introduce

$$r_{HMM,D} = \left\| a^{\text{eff},h} - a^{\text{eff}} \right\|_{\mathcal{F}}, \qquad r_{HMM,A} = \varepsilon \left| b^{\text{eff},h} - b^{\text{eff}} \right|, \tag{33}$$

with the numerical effective data $a^{\text{eff},h}$ and $b^{\text{eff},h}$ from (21) and (19), respectively, and a^{eff} and b^{eff} the exact effective data given in (9). Observe that $r_{HMM,A}$ is scaled to be of order $\mathcal{O}(1)$ with respect to ε.

Note that the macro-micro coupling with $W(K_\delta) = H_{per}^1(K_\delta)$ and $\delta/\varepsilon \in \mathbb{N}$ is optimal for periodic data, e.g., see [2, Proposition 14]. Hence, the upscaling error is *free* of any modeling error (due to non-matching boundary conditions or sampling domain size δ) and thus only depends on the micro discretization.

4.1 A priori error estimates for macro spatial error

As first main result, we prove optimal convergence rates with respect to the macro mesh size H for the spatial macro errors.

Theorem 4.1. *Consider the periodic setting* (29). *Assume that the effective data a^{eff} and $b^{\text{eff}} = \varepsilon^{-1} b^*$ defined in* (9) *satisfy*

$$a^{\text{eff}} \xi \cdot \xi \geq \lambda^{\text{eff}} |\xi|^2, \quad \left| a^{\text{eff}} \xi \right| \leq \Lambda^{\text{eff}} |\xi|, \quad 0 < |b^*| \leq \mathcal{B}^{\text{eff}}, \qquad \forall \xi \in \mathbb{R}^d, \tag{34}$$

with $\lambda^{\text{eff}}, \Lambda^{\text{eff}}, \mathcal{B}^{\text{eff}} > 0$, and that the exact solution u^{eff} to the effective problem (5) *satisfies*

$$u^{\text{eff}} \in \mathcal{C}^0([0,T], H^2(\Omega)), \qquad \partial_t u^{\text{eff}} \in \mathcal{C}^0([0,T], L^2(\Omega)). \tag{35}$$

Assume that condition (**M**) *holds. Denote by $u^H: [0,T] \to V^1(\Omega, \mathcal{T}_H)$ the numerical approximation obtained by* (16) *with initial condition $u^H(0) = P_H g$, where P_H is the $L^2(\Omega)$ projection onto $V^1(\Omega, \mathcal{T}_H)$. Assume that for the considered family of meshes $\{\mathcal{T}_H\}_{H>0}$ there exists $\theta_0 > 0$ such that*

$$\sigma(\mathcal{T}_H) = \min \left\{ |b^* \cdot n_e| \,\big|\, e \in \mathcal{E} \text{ with } b^* \cdot n_e \neq 0 \right\} \geq \theta_0, \qquad \text{uniformly as } H \to 0, \tag{36}$$

where n_e denotes a unit normal vector on e.

Then, there exist $\alpha > 1, H_0, R_0 > 0$ independent of ε and $\varepsilon_0 > 0$ such that for $\varepsilon < \varepsilon_0$, $H < H_0$ and $r_{HMM,D} + r_{HMM,A} < R_0$ it holds

$$\left(\int_0^T \|u^H - u^{\mathrm{eff}}\|_D^2 + \frac{1}{\varepsilon} \|u^H - u^{\mathrm{eff}}\|_{A,\tau}^2 \, dt \right)^{1/2}$$

$$\leq C(H + \varepsilon^{-1/2} H^{3/2} + r_{HMM,D} + \varepsilon^{-1/2} r_{HMM,A}) \|u^{\mathrm{eff}}\|_{\mathcal{C}^0([0,T], H^2(\Omega))}$$

$$+ C_\tau \left(\int_0^T \frac{1}{\varepsilon} \|u^H - P_H u^{\mathrm{eff}}\|_{L^2(\Omega_\tau^c)}^2 \, dt \right)^{1/2},$$

where C and C_τ are independent of u^{eff}, H, ε, $r_{HMM,D}$, $r_{HMM,A}$ and only $C_\tau = \mathcal{O}(\tau^{-2})$ depends on τ.

Note that the error of the numerical solution is measured in space using the DG norm $(\|\cdot\|_D^2 + \varepsilon^{-1}\|\cdot\|_{A,\tau}^2)^{1/2}$, where the advection norm $\|\cdot\|_{A,\tau}$ is scaled as usual with $\mathcal{O}(\sqrt{b^{\mathrm{eff}}}) = \mathcal{O}(\varepsilon^{-1/2})$, e.g., see [17]. We thus obtain sharp convergence rates which have the right scaling with respect to ε, i.e., the relevant terms for advection dominated problems ($\varepsilon \ll 1$) all scale as $\mathcal{O}(\varepsilon^{-1/2})$.

The control of the error in the spatial L^2 norm in the advection dominated regime is obtained by introducing a suitable weighting function, following the ideas of [17,40]. However, the construction used in [17,40] cannot easily be extended to periodic boundary conditions and adapting it introduces the additional error term on Ω_τ^c. We emphasize that this issue is not related to the multiscale character of the proposed method, but would already be an issue for singlescale advection-diffusion problems with *periodic* boundary conditions. For a discussion of the technical details we refer to Remark 5.1.

Remark 4.2. Let us briefly discuss the hypotheses assumed in Theorem 4.1.

The condition $b^* \neq 0$ from (49) ensures that the macroscopic drift is not zero. Note that for $b^* = 0$ the effective problem (5) is purely diffusive and the analysis of the numerical method (16) would be much simpler. The condition $\varepsilon < \varepsilon_0$ is further needed in the proof of Theorem 4.1 to ensure that the advection norm $\|\cdot\|_{A,\tau}$ (scaling as $\varepsilon^{-1/2}$) dominates the unscaled L^2 norm.

The hypothesis (36) ensures that the advective jump penalization is stable, see [19, Eq. (5)]. The assumption (36) could be avoided if b^{eff} would be used to stabilize the advection form B_A in (17) instead of its approximation $b^{\mathrm{eff},h}$. This is however not feasible in practice, as b^{eff} is generally unknown. For a class of meshes in dimension $d = 2$, condition (36) can be proved, see Lemma 5.9. We emphasize that a similar technical assumption would be needed for DG-FEM with other types of variational crimes, e.g., when the advective jump term for singlescale problems as considered in [19] is approximated by numerical quadrature rules. While results about the effect of numerical integration on DG-FEM for advection-diffusion problems (with nonlinear advection) have been derived in [52,53], those results cannot be easily integrated into our analysis, as weaker norms have been used.

4.2 Fully discrete analysis of spatial macro and micro errors

In this section, we give explicit estimates of the HMM errors $r_{HMM,D}$ and $r_{HMM,A}$ defined in (33) in terms of the micro mesh size h.

Exact micro problems. We introduce the exact counterpart to the numerical solutions ρ^h and $\psi^{i,h}$ to (18) and (20), respectively. Let $\bar{\rho} \in H^1_{per}(K_\delta)$ with $\int_{K_\delta} \bar{\rho} \, dx = |K_\delta|$ satisfy the adjoint micro problem

$$\int_{K_\delta} a^\varepsilon \nabla z \cdot \nabla \bar{\rho} \, dx + \int_{K_\delta} b^\varepsilon \cdot \nabla z \, \bar{\rho} \, dx = 0, \qquad \forall z \in H^1_{per}(K_\delta), \tag{37}$$

and, for $1 \leq i \leq d$, let $\bar{\psi}^i \in H^1_{per}(K_\delta)$ be the solution with $\int_{K_\delta} \bar{\psi}^i dx = 0$ to

$$\int_{K_\delta} a^\varepsilon (e_i + \nabla \bar{\psi}^i) \cdot \nabla z \, dx + \int_{K_\delta} b^\varepsilon \cdot (e_i + \nabla \bar{\psi}^i) z \, dx = \int_{K_\delta} \bar{b}^{\mathrm{eff}} \cdot e_i \, z \, dx, \tag{38}$$
$$\forall z \in H^1_{per}(K_\delta),$$

where $\bar{b}^{\mathrm{eff}} = \frac{1}{|K_\delta|} \int_{K_\delta} (a^\varepsilon)^T \nabla \bar{\rho} \, dx + \frac{1}{|K_\delta|} \int_{K_\delta} b^\varepsilon \bar{\rho} \, dx$.

Note that in the periodic setting (29) we have that

$$\bar{\rho}(x) = \rho(\tfrac{x}{\varepsilon}), \qquad \bar{\psi}^i(x) = \varepsilon \chi^i(\tfrac{x}{\varepsilon}), \qquad \text{on } K_\delta, \ i \in \{1, \dots, d\}, \tag{39}$$

11

where ρ and χ^i solve the homogenization cell problems (7) and (8), respectively. Hence, the solutions $\bar{\rho}$ and $\bar{\psi}^i$ exist and are unique and $\bar{\rho}$ satisfies $\bar{\rho}(x) > 0$.

Auxiliary micro problems. As the micro problems (20) are non-symmetric (due to advection terms and non-symmetric tensors a^ε), to derive not only linear, but *sharp* quadratic convergence rates for the micro error the following auxiliary micro problems are required: for $1 \leq i \leq d$, find $\bar{X}^i \in H^1_{per}(K_\delta)$ with $\int_{K_\delta} \bar{X}^i dx = 0$ such that

$$\int_{K_\delta} a^\varepsilon \nabla z \cdot \nabla \bar{X}^i dx = - \int_{K_\delta} a^\varepsilon \nabla z \cdot e_i \bar{\rho} \, dx, \qquad \forall \, z \in H^1_{per}(K_\delta). \tag{40}$$

Note that (40) is a coercive, purely diffusive problem and existence and uniqueness of the solutions $\bar{X}^i$ follow from the Lax-Milgram theorem. The need for such auxiliary adjoint micro problems is known for purely diffusive linear and nonlinear micro problems with non-symmetric tensors, e.g., see [10,27] and [6].

Regularity assumptions. For the solutions $\bar{\rho}$, $\bar{\psi}^i$ and $\bar{X}^i$ to (37), (38) and (40), respectively, we assume

(H1$^\rho$) $\bar{\rho} \in H^2(K_\delta)$ with $\begin{cases} 0 < \rho_0 \leq \bar{\rho}(x) \leq P_0 \text{ on } \overline{K}_\delta \text{ uniformly in } \varepsilon \text{ and } \delta, \\ |\bar{\rho}|_{H^2(K_\delta)} \leq C\varepsilon^{-2}\sqrt{|K_\delta|}; \end{cases}$

(H1) $\bar{\psi}^i \in H^2(K_\delta)$ with $\left|\bar{\psi}^i\right|_{H^2(K_\delta)} \leq C\varepsilon^{-1}\sqrt{|K_\delta|}$ for $1 \leq i \leq d$;

(H1*) $\bar{X}^i \in H^2(K_\delta)$ with $\left|\bar{X}^i\right|_{H^2(K_\delta)} \leq C\varepsilon^{-1}\sqrt{|K_\delta|}$ for $1 \leq i \leq d$.

In the analysis of purely diffusive micro problems hypotheses similar to **(H1)** and **(H1*)** have been used, see [3, Remark 4]. As discussed in [2, Remark 9], in the periodic setting (29) and with (39) in mind, the hypotheses **(H1$^\rho$)** and **(H1)** can be shown if the homogenization cell problems (8) and (7) on the unscaled domain Y have H^2 regularity.

Finally, we assume for the finescale data a^ε and b^ε that

(A1) $a^\varepsilon \in (W^{1,\infty}(\Omega))^{d \times d}$ and $\left|a^\varepsilon_{ij}\right|_{W^{1,\infty}(\Omega)} \leq C\varepsilon^{-1}$ for $1 \leq i,j \leq d$;

(B1) $b^\varepsilon \in (W^{1,\infty}(\Omega))^d$ and $|b^\varepsilon \cdot e_i|_{W^{1,\infty}(\Omega)} \leq C\varepsilon^{-2}$ for $1 \leq i \leq d$.

Those hypotheses are satisfied in view of the periodicity assumption (29) if the periodic data a and b have the regularity $a \in (W^{1,\infty}(Y))^{d \times d}$ and $b \in (W^{1,\infty}(Y))^d$, respectively.

Theorem 4.3. *Consider the periodic setting (29) and assume the hypotheses of Theorem 4.1. If* **(A1)**, **(B1)**, **(H1$^\rho$)**, **(H1)** *and* **(H1*)** *are satisfied and the micro problems (37) and (38) have H^2 regularity, then there exist $\alpha > 1$, $H_0, \hat{h}_0 > 0$ independent of ε and $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$, $H < H_0$ and $h/\varepsilon < \hat{h}_0$ it holds*

$$\left(\int_0^T \|\|u^H - u^{\mathrm{eff}}\|\|_D^2 + \frac{1}{\varepsilon}\|\|u^H - u^{\mathrm{eff}}\|\|_{A,\tau}^2 dt \right)^{1/2}$$

$$\leq C\left(H + \varepsilon^{-1/2}H^{3/2} + \left(\frac{h}{\varepsilon}\right)^2 + \varepsilon^{-1/2}\left(\frac{h}{\varepsilon}\right)^2 \right) + C_\tau \left(\int_0^T \frac{1}{\varepsilon}\|u^H - P_H u^{\mathrm{eff}}\|_{L^2(\Omega_\tau^c)}^2 dt \right)^{1/2},$$

where C and C_τ are independent of H, ε and h and only $C_\tau = \mathcal{O}(\tau^{-2})$ depends on τ.

Finally, we obtain robust convergence towards the finescale solution as $\varepsilon \to 0$.

Corollary 4.4. *Assume the hypotheses of Theorem 4.3. Let u^ε be the solution to the finescale problem (1). For ε sufficiently small, the approximation u^H obtained by the multiscale method (16) satisfies*

$$\left\|u^H - u^\varepsilon\right\|_{L^2(0,T;L^2(\Omega_\tau))} \leq C\left(H^{3/2} + r(\varepsilon) + \left(\frac{h}{\varepsilon}\right)^2 \right) + C_\tau \left\|u^H - P_H u^{\mathrm{eff}}\right\|_{L^2(0,T;L^2(\Omega_\tau^c))},$$

where u^{eff} solves the effective problem (5), $r(\varepsilon)$ is the homogenization error $r(\varepsilon) = \|u^{\mathrm{eff}} - u^\varepsilon\|_{L^2(0,T;L^2(\Omega))}$ and the constants C, C_τ are independent of H, ε and h and only $C_\tau = \mathcal{O}(\tau^{-2})$ depends on τ. According to [26, Theorem 4], it further holds that $r(\varepsilon) \to 0$ as $\varepsilon \to 0$ if the data a^ε and b^ε are sufficiently smooth.

5 Proof of the main results

In this section, we prove the a priori error estimates stated in Section 4. Before proving explicit estimates for the macro and micro spatial error in Section 5.3 and 5.4, respectively, we provide important tools and inequalities in Section 5.1 and derive an error propagation formula in Section 5.2.

5.1 Preliminaries

In this section, we discuss several properties of the L^2 projection on the DG space, introduce a weighting function which is crucial for the error analysis, provide important properties of the numerically upscaled data and introduce a singlescale DG bilinear form for the effective equation (5).

Basic estimates. Let $K \in \mathcal{T}_H$, $v \in H^2(K)$ and denote by $P_H v \in \mathcal{P}^1(K)$ its L^2 projection onto $\mathcal{P}^1(K)$, the set of affine polynomials on K. Then we have the standard local approximation estimate, e.g., see [24, Theorem 3.1.4],

$$|v - P_H v|_{H^k(K)} \leq CH^{2-k}|v|_{H^2(K)}, \qquad \text{for } k \in \{0, 1, 2\}. \tag{41}$$

Further to bound functions $v \in H^1(K)$ on an edge $e \in \mathcal{E}$ being an interface of the element $K \in \mathcal{T}_H$ the following trace inequality is fundamental, see [11, Theorem 3.10],

$$\|v\|_{L^2(e)} \leq CH_e^{1/2}\|\nabla v\|_{L^2(K)} + CH_e^{-1/2}\|v\|_{L^2(K)}, \tag{42}$$

where H_e is the diameter of e.

Time derivative of the L^2 projection. Let $v \in \mathcal{C}^0([0,T], L^2(\Omega))$ with $\partial_t v \in \mathcal{C}^0([0,T], L^2(\Omega))$. Consider further the L^2 projection $P_H \colon L^2(\Omega) \to V^1(\Omega, \mathcal{T}_H)$. Then, it holds that

$$\int_\Omega [\partial_t v(t) - \partial_t P_H v(t)] \, w^H dx = 0, \qquad \forall w^H \in V^1(\Omega, \mathcal{T}_H), \tag{43}$$

i.e., $P_H(\partial_t v) = \partial_t(P_H v)$ (the L^2 projection and the time derivative commute).

Weighting function. Recall that for periodic data the effective drift is defined in (9b) as $b^{\text{eff}} = \varepsilon^{-1} b^*$, where b^* is *independent* of ε. Following [17, 40] we introduce the weighting function $\varphi \colon \overline{\Omega} \to \mathbb{R}$ given by

$$\varphi(x) = \omega(x) + \kappa, \quad \text{with } \omega(x) = \exp(-\eta(x)\Psi_\tau(x)) \text{ and } \eta(x) = 2\frac{b^*}{|b^*|} \cdot x, \tag{44}$$

with $\kappa > 0$ and $\Psi_\tau \in C_0^\infty(\Omega)$ a cutoff function satisfying

$$0 \leq \Psi_\tau \leq 1, \text{ on } \Omega, \quad \Psi_\tau \equiv 1, \text{ in } \Omega_\tau, \quad |\Psi_\tau|_{W^{l,\infty}(\Omega_\tau^c)} \leq C\tau^{-l}, \text{ for } l \in \{1, 2\}, \tag{45}$$

where Ω_τ is the subdomain of Ω defined in (30). Note that φ is *independent* of ε and periodic in Ω. We next introduce the constants $\underline{\varphi}^* = C_\omega^{-1} + \kappa$ and $\overline{\varphi}^* = C_\omega + \kappa$ with $C_\omega > 0$ such that

$$\frac{1}{C_\omega} \leq \omega(x) \leq C_\omega, \text{ on } \Omega_\tau, \qquad \|\omega\|_{W^{2,\infty}(\Omega_\tau)} \leq C_\omega. \tag{46}$$

We emphasize that C_ω is *independent* of τ and ε. The crucial property of φ is that

$$-b^* \cdot \nabla\varphi(x) = 2b^* \cdot \frac{b^*}{|b^*|} \exp(-\eta(x)) = 2|b^*| \exp(-\eta(x)) \geq \frac{2|b^*|}{C_\omega}, \quad \text{for } x \in \Omega_\tau, \tag{47}$$

which is used to robustly control the L^2 error in the advection dominated regime.

Remark 5.1. If we set $\Psi_\tau \equiv 1$ in (44) we recover the weighting function φ used in [7, 17, 40] for the analysis of DG methods for advection-diffusion problems with Dirichlet and/or Neumann boundary conditions. In [7, 17, 40], a key feature of φ was that it is single valued on all interfaces $e \in \mathcal{E}$. As we now consider periodic boundary conditions, continuity of φ is hence not only required in the interior of the spatial domain, but as well over the boundaries of Ω. This is achieved by periodizing φ by introducing the cutoff function Ψ_τ. However, the property (47) is lost within the boundary layer of Ω_τ^c and thus the L^2 error on Ω_τ^c cannot be controlled in the advection dominated regime.

Next, we elaborate the explicit dependence of the constants on the parameter τ in the approximation result of [7, Lemma 5.6].

Lemma 5.2. *Let φ be defined in (44) and $v^H \in V^1(\Omega, \mathcal{T}_H)$. Denote by $P_H(\varphi v^H)$ the L^2 projection of φv^H onto $V^1(\Omega, \mathcal{T}_H)$. Then, the following estimates hold*

$$\left|\varphi v^H - P_H(\varphi v^H)\right|_{H^k(\Omega)} \leq CH^{1-k}\|v^H\|_{(\omega,\tau)}, \qquad \text{for } k = 0, 1,$$

$$\left(\sum_{K \in \mathcal{T}_H} \left\|\varphi v^H - P_H(\varphi v^H)\right\|_{L^2(\partial K)}^2\right)^{1/2} \leq CH^{1/2}\|v^H\|_{(\omega,\tau)},$$

where C is independent of H, ε, τ, ω and $\|\cdot\|_{(\omega,\tau)}$ is given by

$$\|v^H\|_{(\omega,\tau)} = \left(C_\omega^2\|v^H\|_{L^2(\Omega_\tau)}^2 + \|\omega\|_{W^{2,\infty}(\Omega_\tau^c)}^2\|v^H\|_{L^2(\Omega_\tau^c)}^2\right)^{1/2}. \tag{48}$$

Properties of the upscaled data. We next show that the numerically upscaled data has similar properties like the exact effective data if the numerical upscaling is sufficiently accurate.

Lemma 5.3. *Consider the periodic setting (29) and assume (34). Let $a^{\mathrm{eff},h}$ and $b^{\mathrm{eff},h}$ be the numerically upscaled data defined in (21) and (19), respectively. Then there exists $R_1 > 0$ independent of ε such that for $r_{HMM,D} + r_{HMM,A} < R_1$ we have*

$$a^{\mathrm{eff},h}\xi \cdot \xi \geq \frac{\lambda^{\mathrm{eff}}}{2}|\xi|^2, \quad \left|a^{\mathrm{eff},h}\xi\right| \leq 2\Lambda^{\mathrm{eff}}|\xi|, \quad \left|b^{\mathrm{eff},h}\right| \leq 2\mathcal{B}^{\mathrm{eff}}\varepsilon^{-1}, \qquad \forall \xi \in \mathbb{R}^d, \tag{49}$$

where $\lambda^{\mathrm{eff}}, \Lambda^{\mathrm{eff}}$ and $\mathcal{B}^{\mathrm{eff}}$ are the constants from (34).

Proof. Let $\xi \in \mathbb{R}^d$, we straightforwardly obtain from (34) that

$$a^{\mathrm{eff},h}\xi \cdot \xi = a^{\mathrm{eff}}\xi \cdot \xi + (a^{\mathrm{eff},h} - a^{\mathrm{eff}})\xi \cdot \xi \geq \left(\lambda^{\mathrm{eff}} - r_{HMM,D}\right)|\xi|^2,$$

$$\left|a^{\mathrm{eff},h}\xi\right| \leq \left|a^{\mathrm{eff}}\xi\right| + \left\|a^{\mathrm{eff},h} - a^{\mathrm{eff}}\right\|_{\mathcal{F}}|\xi| \leq \left(\Lambda^{\mathrm{eff}} + r_{HMM,D}\right)|\xi|,$$

$$\left|b^{\mathrm{eff},h}\right| \leq \left|b^{\mathrm{eff}}\right| + \varepsilon^{-1}\left|\varepsilon b^{\mathrm{eff},h} - \varepsilon b^{\mathrm{eff}}\right| \leq \left(\mathcal{B}^{\mathrm{eff}} + r_{HMM,A}\right)\varepsilon^{-1}.$$

Hence choosing $R_1 = \min\{\lambda^{\mathrm{eff}}/2, \Lambda^{\mathrm{eff}}, \mathcal{B}^{\mathrm{eff}}\}$ concludes the proof. $\qquad\square$

DG forms with exact data. Consider the periodic setting (29) and introduce the DG bilinear form $B_{\mathrm{eff}}(v, w) = B_{D,\mathrm{eff}}(v, w) + B_{A,\mathrm{eff}}(v, w)$ for $v, w \in H^1(\mathcal{T}_H)$ (with $H^1(\mathcal{T}_H)$ defined in (12)) given by

$$\begin{aligned} B_{D,\mathrm{eff}}(v^H, w^H) &= \int_\Omega a^{\mathrm{eff}}\nabla v \cdot \nabla w \, dx - \int_\Gamma \{a^{\mathrm{eff}}\nabla v\} \cdot [\![w]\!] ds + \int_\Gamma \mu_S[\![v]\!] \cdot [\![w]\!] ds, \\ B_{A,\mathrm{eff}}(v^H, w^H) &= \int_\Omega b^{\mathrm{eff}} \cdot \nabla v \, w \, dx - \int_\Gamma b^{\mathrm{eff}} \cdot [\![v]\!]w_{i_0} \, ds, \end{aligned} \tag{50}$$

with a^{eff} and b^{eff} given in (9) and where $\mu_S|_e = \alpha H_e^{-1}\|a^{\mathrm{eff}}\|_{\mathcal{F}}$ (for $e \in \mathcal{E}$, $\alpha > 1$) and w_{i_0} is the trace of w taken from the inflow element with respect to the exact effective flow b^{eff}.

5.2 Error propagation formula

Consider the periodic setting (29). Let $u^{\mathrm{eff}} \in E$ (with E given in (3)) be the solution of the effective problem (5) with data defined in (9) and u^H be the numerical solution obtained by the multiscale method (16). Further, assume that $u^{\mathrm{eff}}, \partial_t u^{\mathrm{eff}} \in \mathcal{C}^0([0,T], L^2(\Omega))$.

Let $e^H(t) = u^H(t) - P_H u^{\mathrm{eff}}(t)$ where P_H denotes the L^2 projection onto $V^1(\Omega, \mathcal{T}_H)$. For $w^H \in V^1(\Omega, \mathcal{T}_H)$ and $t \in (0, T)$, we then have the error propagation formula

$$\int_\Omega \partial_t e^H(t)w^H dx + B(e^H(t), w^H) = -\int_\Omega \partial_t P_H u^{\mathrm{eff}}(t)w^H dx - B(P_H u^{\mathrm{eff}}(t), w^H) \tag{51a}$$

$$= \int_\Omega [\partial_t u^{\mathrm{eff}}(t) - \partial_t P_H u^{\mathrm{eff}}(t)]w^H dx + B_{\mathrm{eff}}(u^{\mathrm{eff}}(t), w^H) - B(P_H u^{\mathrm{eff}}(t), w^H)$$

$$= B_{\mathrm{eff}}(u^{\mathrm{eff}}(t) - P_H u^{\mathrm{eff}}(t), w^H) \tag{51b}$$

$$+ B_{\mathrm{eff}}(P_H u^{\mathrm{eff}}(t), w^H) - B(P_H u^{\mathrm{eff}}(t), w^H), \tag{51c}$$

where we used the definition (16) of u^H, the consistency of the DG-FEM bilinear form (50) and the property (43) of the L^2 projection.

Let us briefly outline the subsequent analysis. Recall that for the analysis the terms of the bilinear forms B_D and B_A involving the quadrature rule (14) will be written in integral form, see Remark 3.2.

In Section 5.3, we prove the convergence result for the macro error stated in Theorem 4.1. First, in Lemma 5.4 and 5.5, we derive lower bounds for $B(e^H(t), w^H)$ with $w^H = P_H(\varphi e^H(t))$ where φ is the weighting function introduced in (44). Second, in Lemma 5.6, the macro spatial error term (51b) is explicitly bounded with respect to the macro mesh size H and the upscaling error (51c) is estimated in Lemma 5.7 using the abstract HMM error $r_{HMM,.}$ defined in (33). Finally, in Section 5.4 we derive explicit estimates of $r_{HMM,.}$ in terms of the micro mesh size h and prove Theorem 4.3.

5.3 Spatial macro error

In this section, we prove Theorem 4.1. To derive space-discrete a priori error estimates for parabolic problems coercivity of the spatial bilinear form is usually needed in the proofs. However, even in the periodic setting (29), the DG bilinear form B defined in (17) is not coercive with respect to the DG norm $(\|\|\cdot\|\|_D^2 + \varepsilon^{-1}\|\|\cdot\|\|_{A,\tau}^2)^{1/2}$ as the exact effective drift is constant, i.e., $-\operatorname{div} b^{\text{eff}} = 0$, and thus the advective part of the DG norm cannot be controlled robustly in the advection dominated regime. We use the weighting function φ defined in (44) and techniques which are similar to the ones that allowed in [7] to prove an inf-sup condition under the assumption of a relaxed coercivity condition, see [7, Section 5.1.1] for a detailed discussion.

Let us briefly outline the major differences compared to the analysis given in [7]. First, an inf-sup condition like in [7, Theorem 4.3] cannot be expected as already the continuous spatial differential operator only satisfies a Gårding inequality, see (4). Second, the weighting function φ has to be periodized and thus looses its crucial property within a boundary layer. As consequence, the spatial error in the L^2 norm within the boundary layer cannot be controlled independently of ε^{-1}. Third, the advective jump norm $|\cdot|_{*,A}$ introduced in (32) uses the numerically upscaled vector $b^{\text{eff},h}$ as weight instead of the exact effective velocity b^{eff}. On one hand, this allows to derive bounds for the effect of the variational crimes that are more robust with respect to the Péclet number, but, on the other, necessitates the stability assumption (36) to obtain optimal convergence rates.

Lower bounds for the modified macro map. In the following two Lemmas we provide a lower bound for $B(v^H, P_H(\varphi v^H))$ where B is the modified macro map (17), φ the weighting function defined in (44) and P_H the L^2 projection onto $V^1(\Omega, \mathcal{T}_H)$. For doing so, we derive a lower bound for $B(v^H, \varphi v^H)$ and an upper bound for $B(v^H, \varphi v^H - P_H(\varphi v^H))$, respectively.

Lemma 5.4. *Consider the periodic setting* (29). *Let φ be the weighting function given in* (44) *and $v^H \in V^1(\Omega, \mathcal{T}_H)$. If $a^{\text{eff},h}$ and $b^{\text{eff},h}$ satisfy* (49), *then it holds*

$$B_D(v^H, \varphi v^H) \geq \left(\frac{\lambda^{\text{eff}}}{2}\underline{\varphi}^* - C_g \alpha^{-1/2}\overline{\varphi}^* - \Lambda^{\text{eff}}\right)\|\|v^H\|\|_D^2 - \Lambda^{\text{eff}}\|v^H\|_{(\omega,\tau)}^2,$$

$$B_A(v^H, \varphi v^H) \geq \left(\frac{|b^*|}{C_\omega} - \frac{1}{2}C_\omega r_{HMM,A}\right)\frac{1}{\varepsilon}\|v^H\|_{L^2(\Omega_\tau)}^2 + \frac{1}{2C_\omega}\frac{1}{\varepsilon}|v^H|_{*,A}^2$$
$$- \mathcal{B}^{\text{eff}}|\omega|_{W^{1,\infty}(\Omega_\tau^c)}\frac{1}{\varepsilon}\|v^H\|_{L^2(\Omega_\tau^c)}^2,$$

$$\|\|\varphi v^H\|\|_D \leq \overline{\varphi}^*\|\|v^H\|\|_D + \|v^H\|_{(\omega,\tau)}, \quad \|\varphi v^H\|_{L^2(\Omega)} \leq \overline{\varphi}^*\|v^H\|_{L^2(\Omega)}, \quad |\varphi v^H|_{*,A} \leq \overline{\varphi}^*|v^H|_{*,A}.$$

Proof. The lower bound for B_D defined in (17) is obtained following the lines of the proof of [7, Lemma 5.8]. We therefore successively consider the three terms of B_D. For the first term we get

$$\int_\Omega a^{\text{eff},h}\nabla v^H \cdot \nabla(\varphi v^H)dx \geq \underline{\varphi}^*\frac{\lambda^{\text{eff}}}{2}\|\nabla v^H\|_{L^2(\Omega)}^2 - 2\Lambda^{\text{eff}}\|\nabla v^H\|_{L^2(\Omega)}\|v^H\|_{(\omega,\tau)}, \tag{52}$$

where we used the ellipticity and the boundedness (49) of $a^{\text{eff},h}$, the estimates (46) and the definition (48). For the second term of B_D, analogously to [4, Lemma 4.3], using (49) and the continuity and periodicity of φ yields that there exists a constant C_g independent of H and α such that

$$\left|\int_\Gamma \{a^{\text{eff},h}\nabla v^H\} \cdot [\![\varphi v^H]\!]ds\right| \leq C_g\overline{\varphi}^*\alpha^{-1/2}\|\nabla v^H\|_{L^2(\Omega)}|v^H|_{*,D}. \tag{53}$$

For the third term of B_D, we get from the ellipticity of $a^{\text{eff},h}$ and the continuity and periodicity of φ that

$$\int_\Gamma \mu[\![v^H]\!] \cdot [\![\varphi v^H]\!]ds \geq \underline{\varphi}^*\frac{\lambda^{\text{eff}}}{2}|v^H|_{*,D}^2. \tag{54}$$

15

Combining estimates (52)–(54) and using Young's inequality yields the lower bound for B_D.

In contrast, the bound for B_A is different from the one obtained in [7, Lemma 5.8]. Using integration by parts for the first term of B_A and the continuity and periodicity of φ we get

$$
\begin{aligned}
B_A(v^H, \varphi v^H) = & -\frac{1}{2} \sum_{K \in \mathcal{T}_H} \int_K b^{\mathrm{eff},h} \cdot \nabla \varphi (v^H)^2 dx \\
& + \frac{1}{2} \sum_{K \in \mathcal{T}_H} \int_{\partial K} b^{\mathrm{eff},h} \cdot n_K (v^H)^2 \varphi \, ds - \int_\Gamma b^{\mathrm{eff},h} \cdot \llbracket v^H \rrbracket v_i^H \, \varphi \, ds \\
= & -\frac{1}{2\varepsilon} \int_{\Omega_\tau} \varepsilon b^{\mathrm{eff}} \cdot \nabla \varphi (v^H)^2 dx - \frac{1}{2\varepsilon} \int_{\Omega_\tau} (\varepsilon b^{\mathrm{eff},h} - \varepsilon b^{\mathrm{eff}}) \cdot \nabla \varphi (v^H)^2 dx \\
& -\frac{1}{2} \int_{\Omega_\tau^c} b^{\mathrm{eff},h} \cdot \nabla \varphi (v^H)^2 dx + \int_\Gamma b^{\mathrm{eff},h} \cdot \left(\frac{1}{2} \llbracket (v^H)^2 \rrbracket - \llbracket v^H \rrbracket v_i^H \right) \varphi \, ds.
\end{aligned}
$$

Using (47) for φ in Ω_τ, (46) and the identity $1/2 \llbracket (v^H)^2 \rrbracket - \llbracket v^H \rrbracket v_i^H = -1/2 |\llbracket v^H \rrbracket|^2 \cdot n_i$ then yields

$$
\begin{aligned}
B_A(v^H, \varphi v^H) \geq & \left(\frac{|b^*|}{C_\omega} - \frac{1}{2} C_\omega r_{HMM,A} \right) \varepsilon^{-1} \|v^H\|_{L^2(\Omega_\tau)}^2 + \frac{1}{2 C_\omega} \varepsilon^{-1} |v^H|_{*,A}^2 \\
& - \mathcal{B}^{\mathrm{eff}} |\omega|_{W^{1,\infty}(\Omega_\tau^c)} \varepsilon^{-1} \|v^H\|_{L^2(\Omega_\tau^c)}^2.
\end{aligned}
$$

Further, the estimates for φv^H in the (semi)-norms $\|\cdot\|_D$, $\|\cdot\|_{L^2(\Omega)}$ and $|\cdot|_{*,A}$ directly follow from the bounds (46) and the definition (48). $\qquad \square$

We next have the following result.

Lemma 5.5. *Consider the periodic setting* (29). *Let φ be the weighting function given in* (44) *and $v^H \in V^1(\Omega, \mathcal{T}_H)$. Assume that $a^{\mathrm{eff},h}$ and $b^{\mathrm{eff},h}$ satisfy* (49). *Then, we have*

$$
\begin{aligned}
\left| B_D(v^H, \varphi v^H - P_H(\varphi v^H)) \right| &\leq C_D \alpha^{1/2} \|v^H\|_D \|v^H\|_{(\omega, \tau)}, \\
\left| B_A(v^H, \varphi v^H - P_H(\varphi v^H)) \right| &\leq C_A H^{1/2} \varepsilon^{-1} |v^H|_{*,A} \|v^H\|_{(\omega, \tau)},
\end{aligned}
$$

where the constants C_D and C_A are independent of H, ε, α and κ.

Proof. The upper bound for B_D is obtained following the lines of the proof [7, Lemma 5.9] and using the adapted approximation result of Lemma 5.2.

As the jump norm $|\cdot|_{*,A}$ is based on the numerical drift $b^{\mathrm{eff},h}$ (and not on the exact drift like in [7]) the estimate for B_A is different compared to [7, Lemma 4.9] as the influence of the variational crimes $r_{HMM,A}$ do not have to be controlled. Using the definition of the L^2 projection and (49) we get

$$
\begin{aligned}
\left| B_A(v^H, \varphi v^H - P_H(\varphi v^H)) \right| &= \left| \int_\Gamma b^{\mathrm{eff},h} \cdot \llbracket v^H \rrbracket (\varphi v^H - P_H(\varphi v^H))_i ds \right| \\
&\leq \varepsilon^{-1} \int_\Gamma \left| \varepsilon b^{\mathrm{eff},h} \cdot n \right| \left| \llbracket v^H \rrbracket \right| \left| (\varphi v^H - P_H(\varphi v^H))_i \right| ds \\
&\leq \sqrt{2 \mathcal{B}^{\mathrm{eff}}} \varepsilon^{-1} |v^H|_{*,A} \left(\sum_{e \in \mathcal{E}} \|(\varphi v^H - P_H(\varphi v^H))_i\|_{L^2(e)}^2 \right)^{1/2} \\
&\leq C H^{1/2} \varepsilon^{-1} |v^H|_{*,A} \|v^H\|_{(\omega, \tau)},
\end{aligned}
$$

where we used Lemma 5.2. $\qquad \square$

Estimates for error terms (51b) **and** (51c). In the next lemma we derive explicit bounds for the spatial macro error.

Lemma 5.6. *Let the effective data a^{eff} and b^{eff} satisfy* (49) *and assume that the stability condition* (36) *is satisfied. Let $v \in H^2(\Omega)$ and denote by $P_H v$ its L^2 projection onto $V^1(\Omega, \mathcal{T}_H)$. Then, there exists $R_2 > 0$ such that for $r_{HMM,A} < R_2$*

$$
\begin{aligned}
\left| B_{D,\mathrm{eff}}(v - P_H v, w^H) \right| &\leq C H \|v\|_{H^2(\Omega)} \|w^H\|_D, \\
\left| B_{A,\mathrm{eff}}(v - P_H v, w^H) \right| &\leq C \varepsilon^{-1} H^{3/2} \|v\|_{H^2(\Omega)} |w^H|_{*,A},
\end{aligned}
$$

for every $w^H \in V^1(\Omega, \mathcal{T}_H)$ and where C and R_2 are independent of H and ε.

Proof. The estimate for $B_{D,\text{eff}}$ has been derived in [17, Theorem 5.1]. For the advective part we instead follow the arguments of [19, Section 5]. We first observe that for $w, z \in H^1(\mathcal{T}_H)$ defined in (12) the identity $[\![zw]\!] - [\![z]\!]w_{i_0} = [\![w]\!]z_{o_0}$ holds (where i_0 and o_0 denote the inflow and outflow trace, respectively, with respect to b^{eff}). Combining that with integration by parts for the first term of $B_{A,\text{eff}}$ and the definition of the L^2 projection we get

$$
\begin{aligned}
B_{A,\text{eff}}&(v - P_H v, w^H) \\
&= -\int_\Omega b^{\text{eff}} \cdot \nabla w^H(v - P_H v)dx + \int_\Gamma b^{\text{eff}} \cdot \left([\![(v - P_H v)w^H]\!] - [\![v - P_H v]\!]w_{i_0}^H\right)ds \\
&= \int_\Gamma b^{\text{eff}} \cdot [\![w^H]\!](v - P_H v)_{o_0}ds \\
&\leq \sqrt{\mathcal{B}^{\text{eff}}}\varepsilon^{-1}\left(\int_\Gamma |(v - P_H v)_{o_0}|^2 ds\right)^{1/2}\left(\int_\Gamma |\varepsilon b^{\text{eff}} \cdot n||[\![w^H]\!]|^2 ds\right)^{1/2} \\
&\leq C\varepsilon^{-1}H^{3/2}\|v\|_{H^2(\Omega)}\left(\int_\Gamma |\varepsilon b^{\text{eff}} \cdot n||[\![w^H]\!]|^2 ds\right)^{1/2},
\end{aligned}
$$

where we used in the last step the trace inequality (42) and the approximation estimate (41).

To conclude the proof, we have show that

$$
\int_\Gamma |\varepsilon b^{\text{eff}} \cdot n||[\![w^H]\!]|^2 ds \leq C\int_\Gamma |\varepsilon b^{\text{eff},h} \cdot n||[\![w^H]\!]|^2 ds = C|w^H|_{*,A}^2.
$$

Therefore, we prove that for any $e \in \mathcal{E}$ (with associated unit normal vector n) the bound $|\varepsilon b^{\text{eff}} \cdot n| \leq 2|\varepsilon b^{\text{eff},h} \cdot n|$ holds if $r_{HMM,A}$ is sufficiently small. If $b^{\text{eff}} \cdot n = 0$ the bound clearly holds independently of $r_{HMM,A}$. Thus, let $b^{\text{eff}} \cdot n \neq 0$. We then get

$$
\begin{aligned}
|\varepsilon b^{\text{eff}} \cdot n| &\leq |\varepsilon b^{\text{eff},h} \cdot n| + r_{HMM,A} = |\varepsilon b^{\text{eff},h} \cdot n| + \frac{r_{HMM,A}}{|b^* \cdot n|}|\varepsilon b^{\text{eff}} \cdot n| \\
&\leq |\varepsilon b^{\text{eff},h} \cdot n| + \theta_0^{-1}r_{HMM,A}|\varepsilon b^{\text{eff}} \cdot n|,
\end{aligned}
$$

where we used $\varepsilon b^{\text{eff}} = b^*$ (see (9)) and (36). Hence, for $R_2 \leq \theta_0/2$ we have $|\varepsilon b^{\text{eff}} \cdot n| \leq 2|\varepsilon b^{\text{eff},h} \cdot n|$. $\square$

Lemma 5.7. *Let $v \in H^2(\Omega)$ and denote by $P_H v$ its L^2 projection onto $V^1(\Omega, \mathcal{T}_H)$. Then for every $w^H \in V^1(\Omega, \mathcal{T}_H)$ we have*

$$
\begin{aligned}
\left|B_{D,\text{eff}}(P_H v, w^H) - B_D(P_H v, w^H)\right| &\leq Cr_{HMM,D}\|v\|_{H^2(\Omega)}\|w^H\|_D, \\
\left|B_{A,\text{eff}}(P_H v, w^H) - B_A(P_H v, w^H)\right| &\leq C\varepsilon^{-1}r_{HMM,A}\|v\|_{H^2(\Omega)}\|w^H\|_{L^2(\Omega)},
\end{aligned}
$$

with $r_{HMM,\cdot}$ defined in (33) and C independent of H, ε and $r_{HMM,\cdot}$.

Proof. The propagation of the HMM error from the micro scale up to the macro scale has been studied in [7, Lemmas 5.7, 5.12]. In particular, it holds that

$$
\begin{aligned}
\left|B_{D,\text{eff}}(P_H v, w^H) - B_D(P_H v, w^H)\right| &\leq Cr_{HMM,D}\left(\|\nabla P_H v\|_{L^2(\Omega)}^2 + |P_H v|_{*,D}^2\right)^{1/2}\|w^H\|_D, \\
\left|B_{A,\text{eff}}(P_H v, w^H) - B_A(P_H v, w^H)\right| &\leq C\varepsilon^{-1}r_{HMM,A}\left(\|\nabla P_H v\|_{L^2(\Omega)}^2 + |P_H v|_{*,D}^2\right)^{1/2}\|w^H\|_{L^2(\Omega)},
\end{aligned}
$$

from where the result is obtained by using the trace inequality (42) and the estimate (41). $\square$

Proof of macro error estimate. After a preliminary result derived in Lemma 5.8, where a detailed inequality for the evolution of the numerical error is given, we prove Theorem 4.1.

Lemma 5.8. *Consider the periodic setting (29) and assume the hypotheses of Theorem 4.1. Then, there exists $\alpha > 1$, $H_0, R_0, C^* > 0$ independent of ε such that for all $H < H_0$, $r_{HMM,D} + r_{HMM,A} < R_0$ and $t \in (0,T)$ we have*

$$
\begin{aligned}
\partial_t \left\|\varphi^{1/2}e^H(t)\right\|_{L^2(\Omega)}^2 &+ C^*\left(\|e^H(t)\|_D^2 + \frac{1}{\varepsilon}\|e^H(t)\|_{A,\tau}^2\right) \\
&\leq C(H + \varepsilon^{-1/2}H^{3/2} + r_{HMM,D} + \varepsilon^{-1/2}r_{HMM,A})^2\left\|u^{\text{eff}}\right\|_{\mathcal{C}^0([0,T],H^2(\Omega))}^2 \\
&\quad + \bar{C}^2\left\|e^H(t)\right\|_{L^2(\Omega_\tau)}^2 + C_\tau^2\varepsilon^{-1}\left\|e^H(t)\right\|_{L^2(\Omega_\tau^c)}^2,
\end{aligned}
$$

17

where the constants C^*, C, $\bar{C}$ and C_τ are independent of H, ε, $r_{HMM,D}$, $r_{HMM,A}$ and only $C_\tau = \mathcal{O}(\tau^{-2})$ depends on τ.

Proof. Let $t \in (0, T)$ and set $w^H = P_H(\varphi e^H(t))$ in the error propagation formula (51). For readability we write $e^H = e^H(t)$ in what follows.

Starting with the first term of the left-hand side of (51) the definition of the L^2 projection yields

$$\int_\Omega (\partial_t e^H) P_H(\varphi e^H) dx = \int_\Omega (\partial_t e^H) \varphi e^H dx = \frac{1}{2} \int_\Omega \partial_t(\varphi(e^H)^2) dx = \frac{1}{2} \partial_t \left\| \varphi^{1/2} e^H \right\|^2_{L^2(\Omega)}, \qquad (55)$$

as φ defined in (44) is independent of t and strictly positive. Next, let $R_1 > 0$ be given in Lemma 5.3 and assume that $r_{HMM,D} + r_{HMM,A} < R_1$, i.e., $a^{\text{eff},h}$ and $b^{\text{eff},h}$ satisfy (49). Combining Lemmas 5.4, 5.5 and Young's inequality to bound the second term of the left-hand side of (51) yields

$$B(e^H, P_H e^H) \geq \left(\frac{\lambda^{\text{eff}}}{2} \underline{\varphi}^* - C_g \alpha^{-1/2} \overline{\varphi}^* - \Lambda^{\text{eff}} - \frac{1}{2} C_D \alpha^{1/2} \right) \|e^H\|^2_D \qquad (56a)$$

$$+ \left(\frac{|b^*|}{C_\omega} - \frac{1}{2} C_\omega r_{HMM,A} - \frac{1}{2} C_A C_\omega^2 H^{1/2} \right) \varepsilon^{-1} \|e^H\|^2_{L^2(\Omega_\tau)} \qquad (56b)$$

$$+ \left(\frac{1}{2C_\omega} - \frac{1}{2} C_A H^{1/2} \right) \varepsilon^{-1} |e^H|^2_{*,A} \qquad (56c)$$

$$- \left(\Lambda^{\text{eff}} + \frac{1}{2} C_D \alpha^{1/2} \right) \|e^H\|^2_{(\omega,\tau)} \qquad (56d)$$

$$- \left(\mathcal{B}^{\text{eff}} |\omega|_{W^{1,\infty}(\Omega_\tau^c)} + \frac{1}{2} C_A H^{1/2} \|\omega\|^2_{W^{2,\infty}(\Omega_\tau^c)} \right) \varepsilon^{-1} \|e^H\|^2_{L^2(\Omega_\tau^c)}. \qquad (56e)$$

Similar to [7, Theorem 4.3], we then seek for constants $\alpha > 1$, $H_0 > 0$, $R_0 > 0$ such that for an appropriate value of $\kappa > 0$ we have: for all $H < H_0$ and $r_{HMM,D} + r_{HMM,A} < R_0$ the terms (56a)–(56c) can be bounded from below by $C^*(\|e^H\|^2_D + \varepsilon^{-1} \|e^H\|^2_{A,\tau})$ where $C^* > 0$ is a constant independent of H, ε and $r_{HMM,\cdot}$.

We first fix $\alpha > 1$ such that

$$\mathtt{A}(\alpha): \qquad \frac{\lambda^{\text{eff}}}{4} > C_g \alpha^{-1/2}.$$

The parameter $\kappa > 0$ in the definition (44) of φ is next chosen to satisfy

$$\mathtt{B}(\alpha, \kappa): \qquad \frac{4}{3} \underline{\varphi}^* > \overline{\varphi}^*, \qquad \frac{\lambda^{\text{eff}}}{12} \underline{\varphi}^* > \Lambda^{\text{eff}}, \qquad \frac{\lambda^{\text{eff}}}{24} \underline{\varphi}^* > \frac{1}{2} C_D \alpha^{1/2}.$$

Then, we define the threshold value $H_0 > 0$ such that for all $H < H_0$ we have

$$\mathtt{C}(H): \qquad \frac{|b^*|}{4C_\omega} > \frac{1}{2} C_A C_\omega^2 H^{1/2}, \qquad \frac{1}{4C_\omega} > \frac{1}{2} C_A H^{1/2}.$$

Finally, the value $R_0 > 0$ has to satisfy $R_0 \leq \min\{R_1, R_2\}$ (with R_1 and R_2 from Lemma 5.3 and 5.6) and that for $r_{HMM,A} < R_0$ it holds

$$\mathtt{D}(r_{HMM,A}): \qquad \frac{|b^*|}{C_\omega} > C_\omega r_{HMM,A}.$$

With the values of α, κ, H_0 and R_0 chosen according to $\mathtt{A} - \mathtt{D}$ we indeed obtain the desired lower bound for the terms (56a)–(56c) with constant C^* given by

$$C^* = \min \left\{ \frac{1}{2} C_D \alpha^{1/2}, \frac{|b^*|}{4C_\omega}, \frac{1}{4C_\omega} \right\}.$$

Under those conditions we thus get from (51), (55), (56) as well as Lemmas 5.6 and 5.7

$$\frac{1}{2}\partial_t\|\varphi^{1/2}e^H\|_{L^2(\Omega)}^2 + C^*\Big(\|e^H\|_D^2 + \varepsilon^{-1}\|e^H\|_{A,\tau}^2\Big)$$

$$\leq C\|e^H\|_{(\omega,\tau)}^2 + \Big(\mathcal{B}^{\mathrm{eff}}|\omega|_{W^{1,\infty}(\Omega_\tau^c)} + CH^{1/2}\|\omega\|_{W^{2,\infty}(\Omega_\tau^c)}\Big)\varepsilon^{-1}\|e^H\|_{L^2(\Omega_\tau^c)}^2$$

$$+ C\Big[(H + r_{HMM,D})\|P_H(\varphi e^H)\|_D + C\varepsilon^{-1}H^{3/2}|P_H(\varphi e^H)|_{*,A}$$

$$+ C\varepsilon^{-1}r_{HMM,A}\|P_H(\varphi e^H)\|_{L^2(\Omega)}\Big]\|u^{\mathrm{eff}}(t)\|_{H^2(\Omega)}$$

$$\leq \frac{\eta}{2}\Big((H + \varepsilon^{-1/2}H^{3/2} + r_{HMM,D} + \varepsilon^{-1/2}r_{HMM,A}\|u^{\mathrm{eff}}(t)\|_{H^2(\Omega)}\Big)^2$$

$$+ \frac{1}{\eta}(\overline{\varphi}^*)^2\|e^H\|_D^2 + \frac{2}{\eta}\varepsilon^{-1}\Big(((\overline{\varphi}^*)^2 + CH_0C_\omega^2)\|e^H\|_{L^2(\Omega_\tau)}^2 + (\overline{\varphi}^*)^2|e^H|_{*,A}^2\Big) \tag{57}$$

$$+ CC_\omega^2\|e^H\|_{L^2(\Omega_\tau)}^2 + C_\tau^2\varepsilon^{-1}\|e^H\|_{L^2(\Omega_\tau^c)}^2,$$

where in the last step we used the bounds from Lemma 5.4 and Young's inequality (with arbitrary parameter $\eta > 0$). Choosing $\eta > 0$ sufficiently large (but *independent* of ε and τ) the term (57) can be subtracted on both sides, which concludes the proof. $\qquad\square$

Proof of Theorem 4.1. Subtracting $\bar{C}^2\|e^H\|_{L^2(\Omega_\tau)}^2$ on both sides of the bound of Lemma 5.8 we get

$$\partial_t\Big\|\varphi^{1/2}e^H(t)\Big\|_{L^2(\Omega)}^2 + C^*\|e^H(t)\|_D^2 + (C^*\varepsilon^{-1} - \bar{C}^2)\|e^H(t)\|_{L^2(\Omega_\tau)}^2 + C^*\varepsilon^{-1}|e^H(t)|_{*,A}^2$$

$$\leq C(H + \varepsilon^{-1/2}H^{3/2} + r_{HMM,D} + \varepsilon^{-1/2}r_{HMM,A})^2\|u^{\mathrm{eff}}\|_{\mathcal{C}^0([0,T],H^2(\Omega))}^2 + C_\tau^2\varepsilon^{-1}\|e^H(t)\|_{L^2(\Omega_\tau^c)}^2. \tag{58}$$

We next fix the value $\varepsilon_0 > 0$ such that for $\varepsilon < \varepsilon_0$ we have $1/2C^*\varepsilon^{-1} > \bar{C}^2$. Hence, for $\varepsilon < \varepsilon_0$, inequality (58) yields

$$\partial_t\Big\|\varphi^{1/2}e^H(t)\Big\|_{L^2(\Omega)}^2 + \frac{C^*}{2}\Big(\|e^H(t)\|_D^2 + \frac{1}{\varepsilon}\|e^H(t)\|_{A,\tau}^2\Big)$$

$$\leq C(H + \varepsilon^{-1/2}H^{3/2} + r_{HMM,D} + \varepsilon^{-1/2}r_{HMM,A})^2\|u^{\mathrm{eff}}\|_{\mathcal{C}^0([0,T],H^2(\Omega))}^2 + C_\tau^2\varepsilon^{-1}\|e^H(t)\|_{L^2(\Omega_\tau^c)}^2. \tag{59}$$

We then integrate (59) from $t = 0$ to $t = T$ and use $e^H(0) = u^H(0) - P_H g = 0$. Note that we do not apply Gronwall's inequality as that would lead to an error constant C depending on τ^{-1} and ε^{-1}. Finally, the proof is concluded by bounding the difference $u^{\mathrm{eff}} - P_H u^{\mathrm{eff}}$ using the trace inequality (42) and the approximation result (41). $\qquad\square$

Discussion of stability condition (36) **for** $d = 2$**.** In what follows, we discuss situations for the spatial dimension $d = 2$ in which the stability condition (36) can be proved.

Let $K \in \mathcal{T}_H$ and consider the usual refinement strategy

$$\tag{60}$$

where three vertices are introduced at the edge-midpoints of K leading to four congruent new triangles.

Lemma 5.9. *Let $d = 2$ and assume that the family of meshes $\{\mathcal{T}_H\}$ has been obtained by uniformly refining an initial mesh $\mathcal{T}_{H_0}$, i.e., applying at each refinement step the refinement strategy (60) for all triangles $K \in \mathcal{T}_H$. Then, the stability condition (36) is satisfied.*

Proof. Note first that for $\sigma(\mathcal{T}_H)$ defined in (36) it holds $\sigma(\mathcal{T}_H) > 0$ for *any* mesh $\mathcal{T}_H$ (independently of how it has been constructed) and in particular this holds for the initial mesh $\mathcal{T}_{H_0}$. However, we have to prove that for meshes $\{\mathcal{T}_H\}$ obtained by uniform refinement of $\mathcal{T}_{H_0}$, the quantity $\sigma(\mathcal{T}_H)$ is *uniformly bounded away* from 0 if $H \to 0$.

Observe that the refinement strategy (60) only introduces new edges that are parallel to some edge of the mesh before refinement. Thus, we have that $\sigma(\mathcal{T}_H) = \sigma(\mathcal{T}_{H_0})$ for all H and $\sigma(\mathcal{T}_H)$ indeed does not tend to 0 as $H \to 0$. $\qquad\square$

The result of Lemma 5.9 remains valid for non-conformal meshes $\mathcal{T}_H$, i.e., meshes with hanging nodes, that are obtained by recursively refining an initial mesh $\mathcal{T}_{H_0}$, but where in each refinement step not necessarily all elements are refined following (60). As for DG methods meshes with hanging nodes are allowed, the stability condition (36) in dimension $d = 2$ is thus valid for a broad class of family of meshes.

5.4 Spatial micro error

In this section, we prove Theorem 4.3 and Corollary 4.4. Therefore, we derive explicit estimates of the upscaling error $r_{HMM,\cdot}$ defined in (33). Recall that in the periodic setting (29) the upscaling error only consists of the error due to the spatial discretization at micro scale.

Auxiliary upscaled data. We first introduce the auxiliary upscaled data $\bar{a}^{\text{eff}}$ and $\bar{b}^{\text{eff}}$ based on the exact micro solutions $\bar{\rho}$ and $\bar{\psi}^i$ solving (37) and (38), respectively,

$$
\begin{aligned}
\bar{b}^{\text{eff}} &= \frac{1}{|K_\delta|} \int_{K_\delta} (a^\varepsilon)^T \nabla \bar{\rho} \, dx + \frac{1}{|K_\delta|} \int_{K_\delta} b^\varepsilon \bar{\rho} \, dx, \\
\bar{a}^{\text{eff}} &= \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon (I + \nabla \bar{\psi}) \bar{\rho} \, dx - \frac{1}{|K_\delta|} \int_{K_\delta} (a^\varepsilon)^T \nabla \bar{\rho} \, \bar{\psi} \, dx + \frac{1}{|K_\delta|} \int_{K_\delta} (\bar{b}^{\text{eff}} - b^\varepsilon) \bar{\psi} \, \bar{\rho} \, dx,
\end{aligned}
\tag{61}
$$

where $\bar{\psi}(x) = (\bar{\psi}^1(x), \ldots, \bar{\psi}^d(x)) \in \mathbb{R}^{1 \times d}$ and $\nabla \bar{\psi}(x) = (\nabla \bar{\psi}^1(x), \ldots, \nabla \bar{\psi}^d(x)) \in \mathbb{R}^{d \times d}$.

Bounds for the exact micro solutions. In the analysis of the micro error, bounds in the L^2 and H^1 norm for the exact micro functions $\bar{\rho}$, $\bar{\psi}^i$ and $\bar{X}^i$ solving (37), (38) and (40), respectively, with explicit dependence on ε and $|K_\delta|$ are required.

We first observe that by transferring the Poincaré-Wirtinger inequality from Y to K_ε we get for $v \in H^1_{per}(K_\varepsilon)$ with $\int_{K_\varepsilon} v(x) dx = 0$ that

$$
\|v\|_{L^2(K_\varepsilon)} \le C_P \varepsilon \|\nabla v\|_{L^2(K_\varepsilon)},
\tag{62}
$$

where C_P is the Poincaré constant on Y, i.e., independent of ε.

Lemma 5.10. *Consider the periodic setting (29) and assume (2) for a^ε and b^ε. Let $\bar{\rho}$, $\bar{\psi}^i$ and $\bar{X}^i$ be the solutions to the exact micro problems (37), (38) and (40), respectively. Then, we have*

$$
\|\nabla \bar{\rho}\|_{L^2(K_\delta)} \le \frac{\mathcal{B}}{\lambda} \varepsilon^{-1} \|\bar{\rho}\|_{L^2(K_\delta)}, \quad \|\nabla \bar{X}^i\| \le \frac{\Lambda}{\lambda} \|\bar{\rho}\|_{L^2(K_\delta)}, \qquad \text{for } 1 \le i \le d.
$$

If additionally the regularity $\bar{\psi}^i \bar{\rho}, (\bar{\psi}^i)^2 \in H^1_{per}(K_\delta)$ holds and $0 < \rho_0 \le \bar{\rho}(x)$ on $\overline{K_\delta}$, where $\rho_0 > 0$ is independent of ε, then

$$
\|\nabla \bar{\psi}^i\|_{L^2(K_\delta)} \le C \left(\|\bar{\rho}\|_{L^2(K_\delta)} + \frac{1}{\sqrt{|K_\delta|}} \|\bar{\rho}\|^2_{L^2(K_\delta)} \right), \qquad \text{for } 1 \le i \le d,
$$

where C is independent of ε and δ.

Proof. The estimate for $\nabla \bar{\rho}$ is the standard energy estimate for second order elliptic PDEs, see [29, Section 6.2.2], obtained by

$$
\lambda \|\nabla \bar{\rho}\|^2_{L^2(K_\delta)} \le \int_{K_\delta} a^\varepsilon \nabla \bar{\rho} \cdot \nabla \bar{\rho} \, dx = - \int_{K_\delta} b^\varepsilon \cdot \nabla \bar{\rho} \, \bar{\rho} \, dx \le \|b^\varepsilon\|_{L^\infty(K_\delta)} \|\nabla \bar{\rho}\|_{L^2(K_\delta)} \|\bar{\rho}\|_{L^2(K_\delta)}
$$

where we used the ellipticity of a^ε and the micro problem (37).

The bound for $\nabla \bar{X}^i$ is standard (as the problem (40) can be studied in the setting of the Lax-Milgram theorem) and is shown by using the ellipticity and boundedness of a^ε and the micro problem (40)

$$
\lambda \|\nabla \bar{X}^i\|^2_{L^2(K_\delta)} \le \int_{K_\delta} a^\varepsilon \nabla \bar{X}^i \cdot \nabla \bar{X}^i dx = - \int_{K_\delta} a^\varepsilon \nabla \bar{X}^i \cdot e_i \bar{\rho} \, dx \le \Lambda \|\nabla \bar{X}^i\|_{L^2(K_\delta)} \|\bar{\rho}\|_{L^2(K_\delta)}.
$$

Finally, we show the bound for $\nabla \bar{\psi}^i$. Let $i \in \{1, \ldots, d\}$ and take $z = \bar{\psi}^i \bar{\rho}$ (which is in $H^1_{per}(K_\delta)$ by assumption) as test function in the micro problem (38). We then get

$$
\begin{aligned}
&\int_{K_\delta} a^\varepsilon \nabla \bar{\psi}^i \cdot \nabla \bar{\psi}^i \, \bar{\rho} \, dx + \int_{K_\delta} a^\varepsilon (\bar{\psi}^i \nabla \bar{\psi}^i) \cdot \nabla \bar{\rho} \, dx + \int_{K_\delta} b^\varepsilon \cdot (\bar{\psi}^i \nabla \bar{\psi}^i) \bar{\rho} \, dx \\
&= - \int_{K_\delta} a^\varepsilon e_i \cdot \nabla \bar{\psi}^i \bar{\rho} \, dx - \int_{K_\delta} a^\varepsilon e_i \cdot \nabla \bar{\rho} \, \bar{\psi}^i dx + \int_{K_\delta} (\bar{b}^{\text{eff}} - b^\varepsilon) \cdot e_i \bar{\psi}^i \bar{\rho} \, dx,
\end{aligned}
\tag{63}
$$

20

where we observe that the sum of the second and the third term on the left-hand side of (63) vanishes as $\bar{\rho}$ solves (37), $\bar{\psi}^i \nabla \bar{\psi}^i = 1/2 \nabla (\bar{\psi}^i)^2$ and $(\bar{\psi}^i)^2 \in H^1_{per}(K_\delta)$ by assumption, i.e., $(\bar{\psi}^i)^2$ is an admissible test function for (37). Using (2), $\bar{\rho}(x) > \rho_0$ and (63) then yields

$$
\lambda \rho_0 \big\| \nabla \bar{\psi}^i \big\|^2_{L^2(K_\delta)} \leq \Lambda \| \nabla \bar{\psi}^i \|_{L^2(K_\delta)} \| \bar{\rho} \|_{L^2(K_\delta)} + \Lambda \| \nabla \bar{\rho} \|_{L^2(K_\delta)} \| \bar{\psi}^i \|_{L^2(K_\delta)} \\
+ \big(|\bar{b}^{\mathrm{eff}}| + \mathcal{B}\varepsilon^{-1} \big) \| \bar{\psi}^i \|_{L^2(K_\delta)} \| \bar{\rho} \|_{L^2(K_\delta)}. \tag{64}
$$

The bound for $\nabla \bar{\rho}$ at hand, we get for the effective drift $\bar{b}^{\mathrm{eff}}$ given in (61) that

$$
|\bar{b}^{\mathrm{eff}}| \leq \frac{\Lambda}{\sqrt{|K_\delta|}} \| \nabla \bar{\rho} \|_{L^2(K_\delta)} + \frac{\mathcal{B}}{\sqrt{|K_\delta|}} \varepsilon^{-1} \| \bar{\rho} \|_{L^2(K_\delta)} \leq \frac{1}{\sqrt{|K_\delta|}} (\Lambda \mathcal{B} \lambda^{-1} + \mathcal{B}) \varepsilon^{-1} \| \bar{\rho} \|_{L^2(K_\delta)}, \tag{65}
$$

and combining that with the Poincaré estimate (62) and bound (64) yields

$$
\lambda \rho_0 \big\| \nabla \bar{\psi}^i \big\|^2_{L^2(K_\delta)} \leq (\Lambda + \Lambda \mathcal{B} \lambda^{-1} C_P) \| \nabla \bar{\psi}^i \|_{L^2(K_\delta)} \| \bar{\rho} \|_{L^2(K_\delta)} \\
+ C_P \left[(\Lambda \mathcal{B} \lambda^{-1} + \mathcal{B}) \frac{1}{\sqrt{|K_\delta|}} \| \bar{\rho} \|_{L^2(K_\delta)} + \mathcal{B} \right] \| \nabla \bar{\psi}^i \|_{L^2(K_\delta)} \| \bar{\rho} \|_{L^2(K_\delta)},
$$

which concludes the proof. $\qquad\square$

The final bounds are summarized in the following result.

Corollary 5.11. *Consider the periodic setting (29) and assume (2) for a^ε and b^ε. Let $\bar{\rho}$, $\bar{\psi}^i$ and $\bar{X}^i$ be the exact solution to the micro problems (37), (38) and (40), respectively. Assume that $0 < \rho_0 \leq \bar{\rho}(x)$ on $\overline{K_\delta}$ with $\rho_0 > 0$ independent of ε.*
If $\bar{\rho}, \bar{\psi}^i \in H^2(K_\delta)$ for $1 \leq i \leq d$ and $\| \bar{\rho} \|_{L^2(K_\delta)} \leq C \sqrt{|K_\delta|}$, then it holds

$$
\| \nabla \bar{\rho} \|_{L^2(K_\delta)} \leq C \varepsilon^{-1} \sqrt{|K_\delta|}, \quad \big\| \nabla \bar{\psi}^i \big\|_{L^2(K_\delta)} \leq C \sqrt{|K_\delta|}, \quad \big\| \bar{\psi}^i \big\|_{L^2(K_\delta)} \leq C \varepsilon \sqrt{|K_\delta|},
$$
$$
\big\| \nabla \bar{X}^i \big\|_{L^2(K_\delta)} \leq C \sqrt{|K_\delta|}, \quad \big\| \bar{X}^i \big\|_{L^2(K_\delta)} \leq C \varepsilon \sqrt{|K_\delta|},
$$

for $1 \leq i \leq d$ and where C is independent of ε.

Proof. Note first that $\bar{\rho}, \bar{\psi}^i \in H^2(K_\delta)$ implies that $\bar{\rho}\bar{\psi}^i, (\bar{\psi}^i)^2 \in H^1_{per}(K_\delta)$ for $1 \leq i \leq d$ due to the Sobolev embedding $H^2(K_\delta) \hookrightarrow W^{1,4}(K_\delta)$ valid for $d \leq 3$. All the bounds are then obtained by combining the estimate (62), the hypothesis on $\| \bar{\rho} \|_{L^2(K_\delta)}$ and Lemma 5.10. $\qquad\square$

Explicit estimates for the micro error. As key ingredient to estimate the micro error committed in the upscaled data $a^{\mathrm{eff},h}$ and $b^{\mathrm{eff},h}$ defined in (21) and (19), respectively, a priori error estimates for the micro solutions ρ^h and $\psi^{i,h}$ solving (18) and (20), respectively, are needed.

Lemma 5.12. *Consider the periodic setting (29) and assume (2) for a^ε and b^ε. Suppose that condition (M) holds, i.e., the numerical solutions ρ^h and $\psi^{i,h}$ to the micro problems (18) and (20), respectively, exist and are unique. Further let $\bar{p}$ and $\bar{\psi}^i$ be the solutions to (37) and (38), respectively.*
If the assumptions (H1) and (H1$^\rho$) hold and the micro problems (37) and (38) have H^2 regularity, then there exists $\hat{h}_0 > 0$ independent of ε such that for all micro mesh sizes h with $h/\varepsilon < \hat{h}_0$ the following FEM estimates hold

$$
\big\| \nabla \rho^h - \nabla \bar{\rho} \big\|_{L^2(K_\delta)} \leq C \frac{h}{\varepsilon^2} \sqrt{|K_\delta|}, \qquad \big\| \rho^h - \bar{\rho} \big\|_{L^2(K_\delta)} \leq C \left(\frac{h}{\varepsilon} \right)^2 \sqrt{|K_\delta|},
$$
$$
\big\| \nabla \psi^{i,h} - \nabla \bar{\psi}^i \big\|_{L^2(K_\delta)} \leq C \frac{h}{\varepsilon} \sqrt{|K_\delta|}, \qquad \big\| \psi^{i,h} - \bar{\psi}^i \big\|_{L^2(K_\delta)} \leq C \frac{h^2}{\varepsilon} \sqrt{|K_\delta|},
$$

for $1 \leq i \leq d$ and where C is independent of h and ε.

The scaling of the bounds (with respect to ε) in Lemma 5.12 is consistent with (39), i.e., $\bar{\rho}(x) = \rho(x/\varepsilon)$ and $\bar{\psi}^i(x) = \varepsilon \chi^i(x/\varepsilon)$, where ρ and χ^i solve the homogenization cell problems (7) and (8), respectively. In particular, this yields $|\bar{\rho}|_{H^2(K_\delta)} = \varepsilon^{-2} \sqrt{|K_\delta|} |\rho|_{H^2(Y)}$ and $|\bar{\psi}^i|_{H^2(K_\delta)} = \varepsilon^{-1} \sqrt{|K_\delta|} |\chi^i|_{H^2(Y)}$.

Proof. For the proof we refer to [22] where the finite element method for the problems (18) and (20) with Neumann boundary conditions has been studied and optimal FE estimates in the L^2 and H^1 norm have been derived. The result for periodic boundary conditions and the explicit dependence of the constants on ε and $|K_\delta|$ can be elaborated by following the lines of their proof. $\qquad\square$

Using those a priori error estimates, we derive the explicit estimates for the micro error.

Lemma 5.13. *Assume the conditions of Lemma 5.12 and additionally that* (**A1**)*,* (**B1**) *and* (**H1**∗) *are satisfied. Let* $a^{\mathrm{eff},h}$ *and* $b^{\mathrm{eff},h}$ *be the numerically upscaled data defined in* (21) *and* (19)*, respectively, and let* $\bar{a}^{\mathrm{eff}}$ *and* $\bar{b}^{\mathrm{eff}}$ *be given by* (61)*. Then, there exists* $\hat{h}_0 > 0$ *independent of* ε *such that for all micro mesh sizes* h *with* $h/\varepsilon < \hat{h}_0$ *we have*

$$\left\| a^{\mathrm{eff},h} - \bar{a}^{\mathrm{eff}} \right\|_{\mathcal{F}} \le C \left(\frac{h}{\varepsilon} \right)^2, \qquad \left| b^{\mathrm{eff},h} - \bar{b}^{\mathrm{eff}} \right| \le C\varepsilon^{-1} \left(\frac{h}{\varepsilon} \right)^2,$$

where C *is independent of* ε *and* h.

Proof. First, we prove the bound for the error in the effective drift. Let $1 \le i \le d$. Using integration by parts we get from (19) and (61) that

$$(b^{\mathrm{eff},h} - \bar{b}^{\mathrm{eff}}) \cdot e_i = \frac{1}{|K_\delta|} \int_{K_\delta} (b^\varepsilon \cdot e_i - \operatorname{div}(a^\varepsilon e_i))(\rho^h - \bar\rho) dx \le C\varepsilon^{-1} \left\| \rho^h - \bar\rho \right\|_{L^2(K_\delta)} \sqrt{|K_\delta|},$$

where we used assumptions (2) and (**A1**) in the last step. Applying Lemma 5.12 then yields the bound.

Next, for $1 \le i, j \le d$, we split the micro error $a^{\mathrm{eff},h} - \bar{a}^{\mathrm{eff}}$ according to

$$(a^{\mathrm{eff},h} - \bar{a}^{\mathrm{eff}})e_j \cdot e_i = I_1 - I_2 + I_3,$$

where the three error terms I_1, I_2 and I_3 are given by

$$I_1 = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(e_j + \nabla\psi^{j,h}) \cdot e_i \rho^h dx - \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(e_j + \nabla\bar\psi^j) \cdot e_i \bar\rho \, dx,$$

$$I_2 = \frac{1}{|K_\delta|} \int_{K_\delta} (a^\varepsilon)^T \nabla\rho^h \cdot e_i \psi^{j,h} dx - \frac{1}{|K_\delta|} \int_{K_\delta} (a^\varepsilon)^T \nabla\bar\rho \cdot e_i \bar\psi^j dx,$$

$$I_3 = \frac{1}{|K_\delta|} \int_{K_\delta} (b^{\mathrm{eff},h} - b^\varepsilon) \cdot e_i \psi^{j,h} \rho^h dx - \frac{1}{|K_\delta|} \int_{K_\delta} (\bar{b}^{\mathrm{eff}} - b^\varepsilon) \cdot e_i \bar\psi^j \bar\rho \, dx.$$

Before estimating those three terms we note that from assumptions (**H1**) and (**H1**$^\rho$) and Lemma 5.12 we get that for $l \in \{0,1\}$ and sufficiently small h/ε it holds

$$\left| \rho^h \right|_{H^l(K_\delta)} \le C\varepsilon^{-l} \sqrt{|K_\delta|}, \quad \left| \psi^{k,h} \right|_{H^l(K_\delta)} \le C\varepsilon^{-(l-1)} \sqrt{|K_\delta|}, \qquad \text{for } 1 \le k \le d. \tag{66}$$

Estimating I_1. We further decompose I_1 as $I_1 = I_1^\rho + I_1^\psi$ with

$$I_1^\rho = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(e_j + \nabla\psi^{j,h}) \cdot e_i (\rho^h - \bar\rho) dx, \qquad I_1^\psi = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(\nabla\psi^{j,h} - \nabla\bar\psi^j) \cdot e_i \bar\rho \, dx.$$

Using Lemma 5.12 and the bounds (66) we get $|I_1^\rho| \le C(h/\varepsilon)^2$. The estimate for I_1^ψ is more involved. We first get from the auxiliary micro problem (40) that

$$I_1^\psi = -\frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(\nabla\psi^{j,h} - \nabla\bar\psi^j) \cdot \nabla\bar{X}^i dx. \tag{67}$$

Next, we observe that the micro problems (20) and (38) yield that for any $z^h \in S^1(K_\delta, \mathcal{T}_h)$

$$0 = \int_{K_\delta} a^\varepsilon(\nabla\psi^{j,h} - \nabla\bar\psi^j) \cdot \nabla z^h dx + \int_{K_\delta} b^\varepsilon \cdot (\nabla\psi^{j,h} - \nabla\bar\psi^j) z^h dx - \int_{K_\delta} (b^{\mathrm{eff},h} - \bar{b}^{\mathrm{eff}}) \cdot e_j z^h dx. \tag{68}$$

Using $z^h = \mathcal{I}_h \bar{X}^i$ in (68) (the nodal interpolant of $\bar{X}^i$ on $\mathcal{T}_h$) and summing (67) and (68) we get

$$\begin{aligned} I_1^\psi = {}& \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon(\nabla\psi^{j,h} - \nabla\bar\psi^j) \cdot (\nabla\mathcal{I}_h\bar{X}^i - \nabla\bar{X}^i) dx \\ & + \frac{1}{|K_\delta|} \int_{K_\delta} b^\varepsilon \cdot (\nabla\psi^{j,h} - \nabla\bar\psi^j)\mathcal{I}_h\bar{X}^i dx - \frac{1}{|K_\delta|} \int_{K_\delta} (b^{\mathrm{eff},h} - \bar{b}^{\mathrm{eff}}) \cdot e_j \, \mathcal{I}_h\bar{X}^i dx. \end{aligned} \tag{69}$$

For the second term on the right-hand side of (69), using an integration by parts and observing that the boundary term cancels (due to the periodicity of the data and the periodic coupling) we obtain

$$\int_{K_\delta} b^\varepsilon \cdot (\nabla \psi^{j,h} - \nabla \bar\psi^j) \mathcal{I}_h \bar X^i dx = -\int_{K_\delta} \mathrm{div}(b^\varepsilon)(\psi^{j,h} - \bar\psi^j)\mathcal{I}_h \bar X^i dx - \int_{K_\delta} b^\varepsilon \cdot \nabla \mathcal{I}_h \bar X^i (\psi^{j,h} - \bar\psi^j)dx. \tag{70}$$

Combining (69) and (70) with assumptions (2) and (**B1**) yields the estimate

$$\left| I_1^\psi \right| \le \frac{C}{|K_\delta|} \Big[\left\| \nabla \psi^{j,h} - \nabla \bar\psi^j \right\|_{L^2(K_\delta)} \left\| \nabla \mathcal{I}_h \bar X^i - \nabla \bar X^i \right\|_{L^2(K_\delta)}$$
$$+ \left(\varepsilon^{-2} \left\| \psi^{j,h} - \bar\psi^j \right\|_{L^2(K_\delta)} + \sqrt{|K_\delta|} \left| b^{\mathrm{eff},h} - \bar b^{\mathrm{eff}} \right| \right) \left\| \mathcal{I}_h \bar X^i \right\|_{L^2(K_\delta)}$$
$$+ \varepsilon^{-1} \left\| \psi^{j,h} - \bar\psi^j \right\|_{L^2(K_\delta)} \left\| \nabla \mathcal{I}_h \bar X^i \right\|_{L^2(K_\delta)} \Big].$$

Observing that similarly to (66) we have that $\|\mathcal{I}_h \bar X^i\|_{H^l(K_\delta)} \le C \varepsilon^{-(l-1)} \sqrt{|K_\delta|}$, for $l \in \{0,1\}$, and combining that with Lemma 5.12, the classical interpolation estimate $\|\nabla \mathcal{I}_h \bar X^i - \nabla \bar X^i\|_{L^2(K_\delta)} \le Ch|\bar X^i|_{H^2(K_\delta)}$ (see [24, Theorem 3.1.6]), (**H1***) and the bound for $b^{\mathrm{eff},h} - \bar b^{\mathrm{eff}}$ then yields $|I_1^\psi| \le C(h/\varepsilon)^2$.

Estimating I_2. We decompose the error I_2 into $I_2 = I_2^\rho + I_2^\psi$ where

$$I_2^\rho = \frac{1}{|K_\delta|}\int_{K_\delta} (a^\varepsilon)^T (\nabla\rho^h - \nabla\bar\rho) \cdot e_i \psi^{j,h} dx, \qquad I_2^\psi = \frac{1}{|K_\delta|}\int_{K_\delta} (a^\varepsilon)^T \nabla\bar\rho \cdot e_i (\psi^{j,h} - \bar\psi^j) dx.$$

Using integration by parts for I_2^ρ, where the boundary term again vanishes due to the periodic data and coupling, we get

$$I_2^\rho = -\frac{1}{|K_\delta|}\int_{K_\delta} \mathrm{div}(a^\varepsilon e_i)(\rho^h - \bar\rho)\psi^{j,h} dx - \frac{1}{|K_\delta|}\int_{K_\delta} a^\varepsilon e_i \cdot \nabla\psi^{j,h}(\rho^h - \bar\rho)dx$$
$$\le \frac{C}{|K_\delta|}\left\| \rho^h - \bar\rho \right\|_{L^2(K_\delta)} \left[\varepsilon^{-1}\left\| \psi^{j,h} \right\|_{L^2(K_\delta)} + \left\| \nabla\psi^{j,h} \right\|_{L^2(K_\delta)} \right] \le C\left(\frac{h}{\varepsilon} \right)^2,$$

where we used (**A1**), Lemma 5.12 and the bounds (66). The bound for I_2^ψ is straightforwardly obtained from Lemma 5.12 and (66) as

$$\left| I_2^\psi \right| \le \frac{C}{|K_\delta|}\|\nabla\bar\rho\|_{L^2(K_\delta)}\left\| \psi^{j,h} - \bar\psi^j \right\|_{L^2(K_\delta)} \le C\left(\frac{h}{\varepsilon} \right)^2.$$

Estimating I_3. To bound the term I_3 we split it into three parts $I_3 = I_3^\rho + I_3^\psi + I_3^b$ with

$$I_3^\rho = \frac{1}{|K_\delta|}\int_{K_\delta} (b^{\mathrm{eff},h} - b^\varepsilon) \cdot e_i \psi^{j,h}(\rho^h - \bar\rho)dx,$$

$$I_3^\psi = \frac{1}{|K_\delta|}\int_{K_\delta} (b^{\mathrm{eff},h} - b^\varepsilon) \cdot e_i (\psi^{j,h} - \bar\psi^j)\bar\rho\, dx, \qquad I_3^b = \frac{1}{|K_\delta|}\int_{K_\delta} (b^{\mathrm{eff},h} - \bar b^{\mathrm{eff}}) \cdot e_i \bar\psi^j \bar\rho\, dx.$$

Using the bound (65), the hypotheses (2), Lemma 5.12, (66) and the estimate for $b^{\mathrm{eff},h} - \bar b^{\mathrm{eff}}$ we get

$$|I_3^\rho| \le \frac{C}{|K_\delta|}\varepsilon^{-1}\left\| \psi^{j,h} \right\|_{L^2(K_\delta)}\left\| \rho^h - \bar\rho \right\|_{L^2(K_\delta)} \le C\left(\frac{h}{\varepsilon} \right)^2,$$

$$\left| I_3^\psi \right| \le \frac{C}{|K_\delta|}\varepsilon^{-1}\left\| \psi^{j,h} - \bar\psi^j \right\|_{L^2(K_\delta)}\|\bar\rho\|_{L^2(K_\delta)} \le C\left(\frac{h}{\varepsilon} \right)^2,$$

$$\left| I_3^b \right| \le \frac{C}{|K_\delta|}\left| b^{\mathrm{eff},h} - \bar b^{\mathrm{eff}} \right|\left\| \bar\psi^j \right\|_{L^2(K_\delta)}\|\bar\rho\|_{L^2(K_\delta)} \le C\left(\frac{h}{\varepsilon} \right)^2. \qquad \Box$$

Proof of Theorem 4.3. Observe that for the effective data $\bar a^{\mathrm{eff}}$ and $\bar b^{\mathrm{eff}}$ defined in (61) the assumptions of the periodic setting (29) directly yield $\bar a^{\mathrm{eff}} = a^{\mathrm{eff}}$ and $\bar b^{\mathrm{eff}} = b^{\mathrm{eff}}$ with a^{eff} and b^{eff} given in (9). Hence, the fully discrete a priori error estimates are obtained by combining Theorem 4.1 and Lemma 5.13. $\qquad \Box$

Proof of Corollary 4.4. We decompose the total error as $\|u^H - u^\varepsilon\|_{L^2(0,T;L^2(\Omega_\tau))} \le \|u^H - u^{\mathrm{eff}}\|_{L^2(0,T;L^2(\Omega_\tau))} + \|u^{\mathrm{eff}} - u^\varepsilon\|_{L^2(0,T;L^2(\Omega_\tau))}$ and estimate the two terms separately. For small values of ε, the terms involving
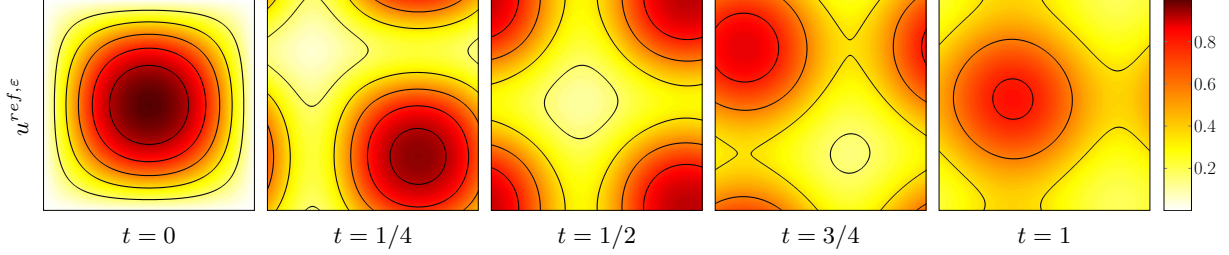
Figure 1: Reference solution $u^{ref,\varepsilon}$ for the effective problem (5) associated to the test problem considered in Section 6.1 with $\varepsilon = 1/10$. Solution plotted on $\Omega = (0,1)^2$ at different times $t \in [0,1]$.

negative powers of ε dominate in the right-hand side of the estimate derived in Theorem 4.3. Thus, we get for ε sufficiently small that

$$\left(\int_0^T \left\| u^H - u^{\text{eff}} \right\|_{L^2(\Omega_\tau)}^2 + \left| u^H - u^{\text{eff}} \right|_{*,A}^2 dt \right)^{1/2}$$

$$\leq C \left(H^{3/2} + \left(\frac{h}{\varepsilon} \right)^2 \right) + C_\tau \left(\int_0^T \left\| u^H - P_H u^{\text{eff}} \right\|_{L^2(\Omega_\tau^c)}^2 dt \right)^{1/2}.$$

Combining that with the homogenization result that $u^\varepsilon \to u^{\text{eff}}$ converges strongly in $L^2(0,T;L^2(\Omega))$ as $\varepsilon \to 0$ (see Section 2) concludes the proof. □

6 Numerical results

In this section, we provide numerical tests corroborating the results of Theorem 4.3 and Corollary 4.4 and we illustrate the applicability of the proposed multiscale method beyond the periodic setting (2).

6.1 Convergence rates of spatial errors

In the periodic setting of Theorem 4.3, we study the convergence rates of the spatial macro and micro errors by comparing the HMM solution u^H to the effective solution u^{eff} in the spatial DG norm.

Let $\Omega = (0,1)^2$ and $T = 1$. Consider (1) with periodic boundary conditions, $g(x) = \sin(\pi x_1)\sin(\pi x_2)$, $f \equiv 0$ and periodic data $a^\varepsilon(x) = a(x/\varepsilon)$, $b^\varepsilon(x) = \varepsilon^{-1}b(x/\varepsilon)$ with

$$a(y) = \frac{1}{100}(2 + \sin(2\pi y_1)\cos(2\pi y_2))Id, \qquad b(y) = \begin{pmatrix} \frac{1}{10}\sin(2\pi y_1)\cos(2\pi y_2) + \frac{1}{10} \\ \frac{1}{10}\cos(2\pi y_1)\sin(2\pi y_2) - \frac{1}{10} \end{pmatrix}, \qquad (71)$$

where $b(y)$ has non-zero mean and is compressible. While the effective model (5) has constant (upscaled) data (9), the effective drift b^{eff} still scales like $1/\varepsilon$. For the convergence tests we thus choose $\varepsilon = 1/50$.

Reference solution for effective problem. As the function u^{eff} solving the effective problem (5) is not known analytically, we compute a reference solution $u^{ref,\varepsilon}$ as follows. First, we precompute a^{eff} and b^* defined in (9) up to a relative error $\mathcal{O}(10^{-7})$ by solving the cell problems (8) and (7) using a standard $\mathcal{P}^1$-FEM on a fine, uniform mesh on Y. We get $a^{\text{eff}} \approx (0.0191, -0.0012; -0.0013, 0.0190)$ and $b^* \approx (0.0850, -0.0972)^T$ (note that $b^* \neq \int_Y b(y)dy$). Second, we solve the effective problem (5) with the precomputed values for a^{eff} and $b^{\text{eff}} = \varepsilon^{-1}b^*$ by using the midpoint Runge-Kutta method (with uniform time step $\Delta t = 1/1024$) in time and the $\mathcal{P}^1$-FEM (on a uniform mesh on Ω with 1024^2 grid points) in space. In Figure 1, we plot $u^{ref,\varepsilon}$ for $\varepsilon = 1/10$.

Convergence rates. Using (16), we compute the HMM solution u^H with sampling domain size $\delta = \varepsilon$. On Ω and K_δ, we use uniform triangular meshes $\mathcal{T}_H$ and $\mathcal{T}_h$ (nodes on opposite faces are identified) with N_{mac} and N_{mic} grid points in each spatial dimension, respectively, i.e., $H \sim N_{mac}^{-1}$, $h/\varepsilon \sim N_{mic}^{-1}$. To integrate (16) in time, we apply the midpoint Runge-Kutta method with uniform time step $\Delta t = 1/1024$, i.e., the same time discretization as for $u^{ref,\varepsilon}$. Hence, the error $u^H - u^{ref,\varepsilon}$ is dominated by spatial errors. We then compare the HMM solution u^H to $u^{ref,\varepsilon}$ by calculating the following measure for the error

$$e_{L^2(DG)} = \left(\int_0^T \| u^H - u^{ref,\varepsilon} \|_D^2 + \frac{1}{\varepsilon} \| u^H - u^{ref,\varepsilon} \|_{A,0}^2 dt \right)^{1/2}, \qquad (72)$$
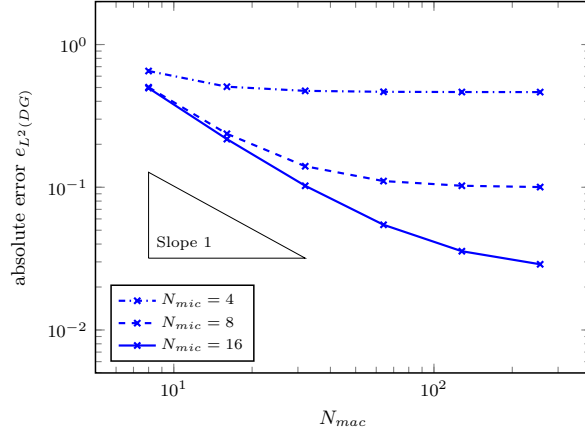
Figure 2: Convergence of the spatial errors of the multiscale method (16). Test problem from Section 6.1 for $\varepsilon = 1/50$. Comparison to the reference solution $u^{ref,\varepsilon}$ introduced in Section 6.1. Absolute error measured by $e_{L^2(DG)}$ defined in (72) in dependence of N_{mac} for $N_{mac} \in \{8, 16, 32, 64, 128, 256\}$. Different error curves correspond to different micro meshes with $N_{mic} \in \{4, 8, 16\}$.

with $\|\cdot\|_D$ and $\|\cdot\|_{A,\tau}$ given in (31). Observe that in contrast to Theorem 4.3, we set $\tau = 0$ in $\|\cdot\|_{A,\tau}$. Further, the integral $\int_0^T \cdot\, dt$ is evaluated using the trapezoidal rule on subintervals of length Δt.

In Figure 2, for $\varepsilon = 1/50$, we plot $e_{L^2(DG)}$ given in (72) with respect to $N_{mac} \sim H^{-1}$ for $N_{mic} = 4, 8, 16$. On one hand, the micro error gets dominant for fine macro meshes (large N_{mac}) as $e_{L^2(DG)}$ saturates at levels depending on N_{mic}, which decrease by a factor 4 when doubling N_{mic}. This confirms the quadratic micro rate derived in Theorem 4.3. On the other hand, when macro error dominates, we get linear convergence in N_{mac} which is optimal for diffusion dominated problems. This suggests that for $\varepsilon = 1/50$ the advection does not yet sufficiently dominate to obtain the superlinear rate predicted in Theorem 4.3 for small ε. Further, we do not observe any pollution due to the term $\varepsilon^{-1/2}\|u^H - P_H u^{\text{eff}}\|_{L^2(\Omega_\tau^c)}$ present in Theorem 4.3.

6.2 Comparison between HMM solution and finescale solution.

For periodic data, we address the convergence in the spatial L^2 norm (as $\varepsilon \to 0$) of the HMM solution u^H to the finescale solution u^ε and, for locally periodic data, we qualitatively study the effects of advection.

Reference solution for finescale problem. For the various tests with $\Omega \times (0, T) = (0, 1)^2 \times (0, 1)$ we approximate u^ε solving (1) by combining the midpoint Runge-Kutta method in time (using $\Delta t = 1/1024$ as time step) with the standard $\mathcal{P}^1$-FEM in space. To resolve the finescale details of a^ε and b^ε, we use a fine uniform triangular mesh on Ω with 1024^2 grid points.

Parameters for multiscale method. For the multiscale scheme (16) we use uniform triangular meshes on Ω and K_δ (for which we set $\delta = \varepsilon$) with $N_{mac} = 256$ and $N_{mic} = 64$ spatial points in each dimension, respectively. For integration in time, we apply the midpoint Runge-Kutta method with time step $\Delta t = 1/1024$. Further, for the tests with locally periodic data, we replace $a^\varepsilon(x)$ and $b^\varepsilon(x)$ in the micro problems (18), (20) and the formulas (19), (21) by $a(x_K, x/\varepsilon)$ and $\varepsilon^{-1}b(x_K, x/\varepsilon)$ where x_K is the macro quadrature point in K_δ (this choice is optimal for locally periodic diffusion problems, see [2, Remark 13]).

Convergence as $\varepsilon \to 0$. We consider (1) with $\Omega = (0, 1)^2$, $T = 1$, periodic boundary conditions, $f \equiv 0$ and $g(x) = \sin(\pi x_1)\sin(\pi x_2)$ for four sets of data (called Test A–D). While Test A and B have periodic data $a^\varepsilon(x) = a(x/\varepsilon)$, $b^\varepsilon(x) = \varepsilon^{-1}b(x/\varepsilon)$, Test C and D have locally periodic data $a^\varepsilon(x) = a(x, x/\varepsilon)$, $b^\varepsilon(x) = \varepsilon^{-1}b(x, x/\varepsilon)$, i.e., they additionally have slow variations at macro scale. We choose

$$
\begin{aligned}
\text{Test A:} \qquad & a(y) = \frac{1}{100}Id, \qquad b(y) = \begin{pmatrix} \sin(2\pi y_1)\cos(2\pi y_2) \\ -\cos(2\pi y_1)\sin(2\pi y_2) \end{pmatrix}, \\[2mm]
\text{Test B:} \qquad & a(y) = \frac{1}{100}Id, \qquad b(y) = \begin{pmatrix} \sin(2\pi y_1) \\ -\frac{1}{10}\cos(2\pi y_1) \end{pmatrix}, \\[2mm]
\text{Test C,D:} \qquad & a(x, y) = \frac{1}{100}(3 + \sin(2\pi(x_1 + x_2)) + \sin(2\pi y_1)\cos(2\pi y_2))Id, \\[2mm]
\text{Test C,D:} \qquad & b(x, y) = \begin{pmatrix} \frac{1}{2}(\sin(2\pi x_1)\cos(2\pi x_2) - \vartheta \sin(2\pi y_1)\cos(2\pi y_2)) \\ \frac{1}{2}(-\sin(2\pi x_1) + \cos(2\pi y_1)\sin(2\pi y_2)) \end{pmatrix},
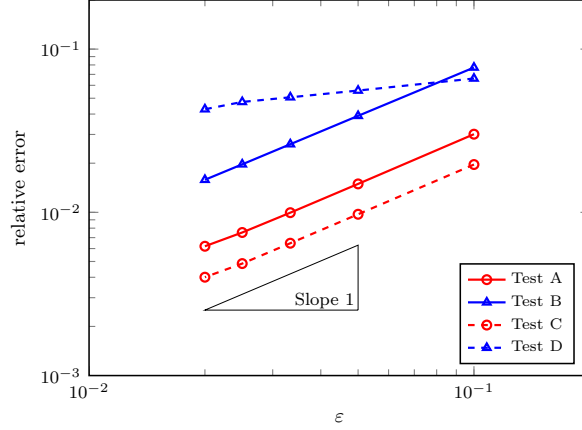\end{aligned}
\tag{73}
$$

25

Figure 3: Error between numerical solution obtained by the multiscale method (16) and the finescale solution u^ε as $\varepsilon \to 0$. Test problems Test A–D defined in (73). Relative error in $L^2(0,T;L^2(\Omega))$ norm.

with $\vartheta = 1$ for Test C, $\vartheta = -1$ for Test D and take $\varepsilon = 1/10, 1/20, 1/30, 1/40, 1/50$.

Test A and B have constant diffusion and flows with $\int_Y b\,dy = 0$, but it holds $\mathrm{div}_y\,b = 0$ for Test A (the usual cellular flow, e.g., see [44, Eq. (62)]), while $\mathrm{div}_y\,b \neq 0$ for Test B. Test C and D have data varying at macro and micro scale (with identical a^ε) and flows with $\int_Y b(x,y)dy \neq 0$ in general. Again the difference is, that $\mathrm{div}_y\,b(x,y) = 0$ for Test C, while $\mathrm{div}_y\,b(x,y) \neq 0$ for Test D. Note that for the periodic Tests A and B, Corollary 4.4 indicates that the HMM solution u^H is close (in the $L^2(0,T;L^2(\Omega))$ norm) to the finescale solution u^ε if ε and the numerical errors are small.

We now study the behavior of the L^2 error as $\varepsilon \to 0$. Therefore, for $\varepsilon = 1/10, \ldots, 1/50$, we compute u^H and the reference solution (FEM) for the Tests A–D and calculate the relative error in the $L^2(0,T;L^2(\Omega))$ norm, which we plot in Figure 3 in dependence of ε. For all tests, the error decreases as ε gets smaller. While for Test A–C the convergence rate is linear, the rate for Test D lies between 0.2 and 0.5. For the advection-diffusion problem (1), explicit rates in ε for the error $\|u^{\mathrm{eff}} - u^\varepsilon\|_{L^2(0,T;L^2(\Omega))}$ have only been proved for periodic data and incompressible flows, i.e., only for Test A (see [47, Theorem 2], where linear convergence has been shown). Interestingly, the method (16) even yields good results for Tests C and D which are beyond the periodic setting (29).

Qualitative study for locally periodic data. We next illustrate the effects of advection on the macro dynamics encountered in Test C and D (see (73)) and show results for yet other families of locally periodic data.

We first consider (1) for $\Omega = (0,1)^2$, $T = 1$, $f \equiv 0$, periodic boundary conditions, initial condition $g(x) = \sin(\pi x_1)\sin(\pi x_2)$ with locally periodic data

$$a^\varepsilon(x) = a(x, \tfrac{x}{\varepsilon}), \quad b^\varepsilon(x) = \varepsilon^{-1}b(x, \tfrac{x}{\varepsilon}), \quad b(x,y) = \varsigma_{mac}b_{mac}(x) + \varsigma_{mic}b^\vartheta_{mic}(y),$$

$$a(x,y) = \frac{1}{100}(3 + \sin(2\pi(x_1 + x_2)) + \sin(2\pi y_1)\cos(2\pi y_2))Id,$$

$$b_{mac}(x) = \frac{1}{2}\begin{pmatrix} \sin(2\pi x_1)\cos(2\pi x_2) \\ -\sin(2\pi x_1) \end{pmatrix}, \quad b^\vartheta_{mic}(y) = \frac{1}{2}\begin{pmatrix} -\vartheta\sin(2\pi y_1)\cos(2\pi y_2) \\ \cos(2\pi y_1)\sin(2\pi y_2) \end{pmatrix},$$

(74)

where choosing $\varsigma_{mac}, \varsigma_{mic} \in \{0,1\}$, $\vartheta \in \{-1,1\}$ yields six test problems (all with the same diffusion tensor) with different characteristics of the advection

- $\varsigma_{mac} = \varsigma_{mic} = 0$: $b^\varepsilon \equiv 0$ (purely diffusive problem), see Figure 4;
- $\varsigma_{mac} = 0, \varsigma_{mic} = 1, \vartheta = 1$: periodic $b^\varepsilon(x) = \varepsilon^{-1}b(x/\varepsilon)$, $\mathrm{div}_y\,b(y) = 0$, $\int_Y b(y)dy = 0$, Figure 5;
- $\varsigma_{mac} = 0, \varsigma_{mic} = 1, \vartheta = -1$: periodic $b^\varepsilon(x) = \varepsilon^{-1}b(x/\varepsilon)$, $\mathrm{div}_y\,b(y) \neq 0$, $\int_Y b(y)dy = 0$, see Figure 6;
- $\varsigma_{mac} = 1, \varsigma_{mic} = 0$: $b^\varepsilon(x) = \varepsilon^{-1}b(x)$ only has variations at macro scale, see Figure 7;
- $\varsigma_{mac} = 1, \varsigma_{mic} = 1, \vartheta = 1$: locally periodic $b^\varepsilon(x) = \varepsilon^{-1}b(x, x/\varepsilon)$, $\mathrm{div}_y\,b(x,y) = 0$, $\int_Y b(x,y)dy \neq 0$, see Figure 8;
- $\varsigma_{mac} = 1, \varsigma_{mic} = 1, \vartheta = -1$: locally periodic $b^\varepsilon(x) = \varepsilon^{-1}b(x, x/\varepsilon)$, $\mathrm{div}_y\,b(x,y) \neq 0$, $\int_Y b(x,y)dy \neq 0$, see Figure 9;

For those test problems, we plot in Figure 4–9 the reference solution (FEM) and the HMM solution (DG-HMM) at times $t = 0, 1/4, 1/2, 3/4, 1$ for the value $\varepsilon = 1/20$.
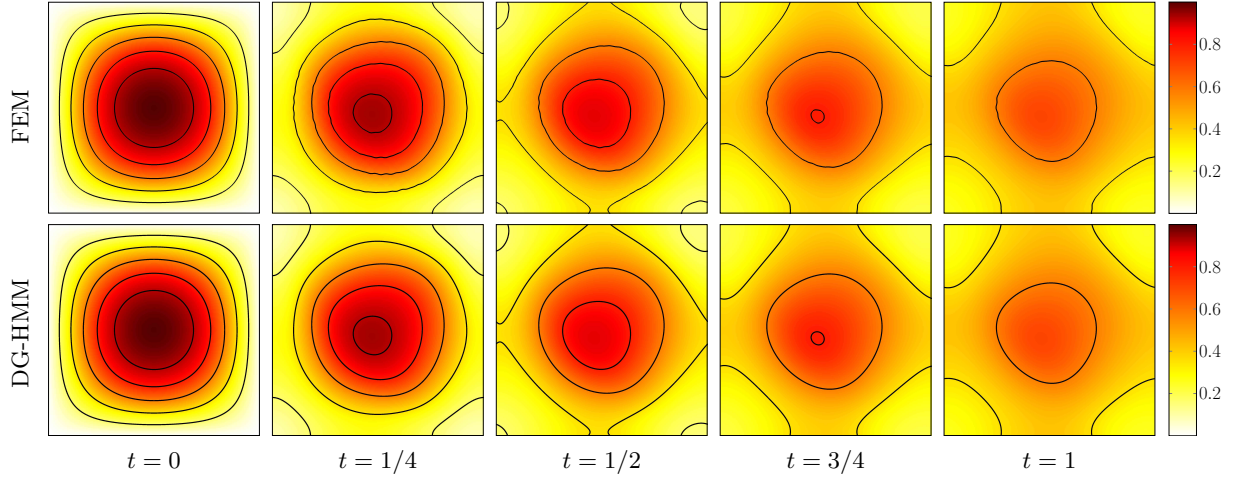
26

Figure 4: Test defined by (74) with $\varepsilon = 1/20$ and $\varsigma_{mac} = \varsigma_{mic} = 0$, i.e., $b^\varepsilon(x) = 0$. FE approximation of u^ε and DG-HMM solution u^H plotted on $\Omega = (0,1)^2$ at different times $t \in [0,T]$.
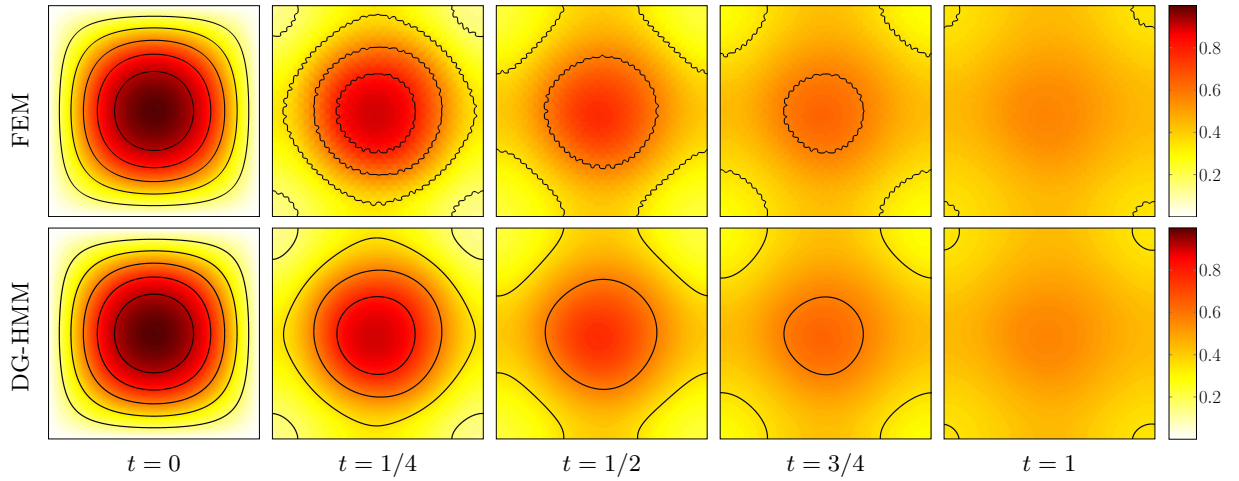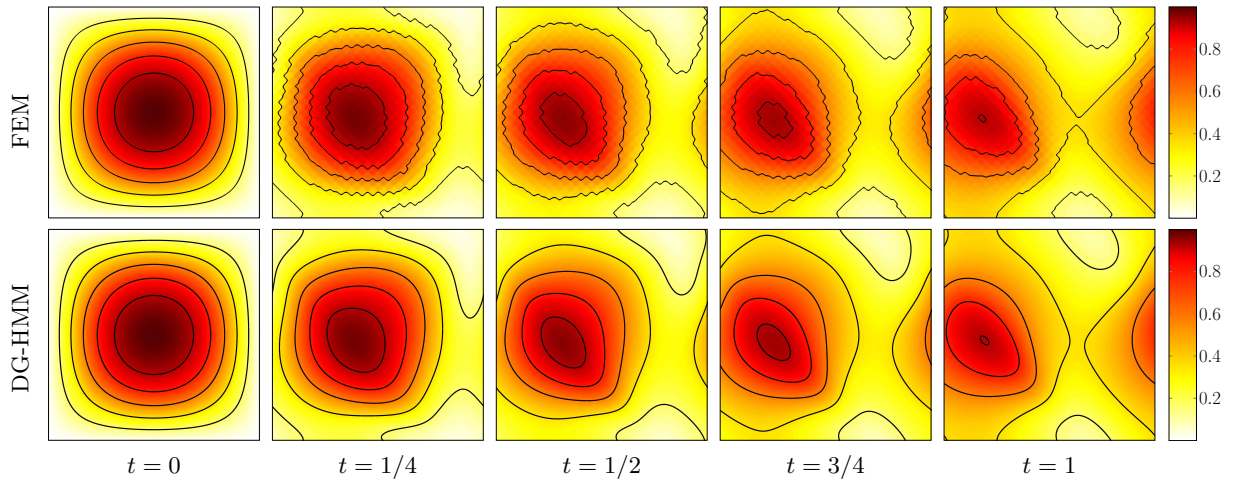

Figure 5: Test defined by (74) with $\varepsilon = 1/20$ and $\varsigma_{mac} = 0$, $\varsigma_{mic} = 1$, $\vartheta = 1$, i.e., $b^\varepsilon(x) = b(\frac{x}{\varepsilon})$ with $\text{div}_y\, b(y) = 0$. FE approximation of u^ε and DG-HMM solution u^H plotted on $\Omega = (0,1)^2$ at different times $t \in [0,T]$.

Figure 6: Test defined by (74) with $\varepsilon = 1/20$ and $\varsigma_{mac} = 0$, $\varsigma_{mic} = 1$, $\vartheta = -1$, i.e., $b^\varepsilon(x) = b(\frac{x}{\varepsilon})$ with $\text{div}_y\, b(y) \neq 0$. FE approximation of u^ε and DG-HMM solution u^H plotted on $\Omega = (0,1)^2$ at different times $t \in [0,T]$.
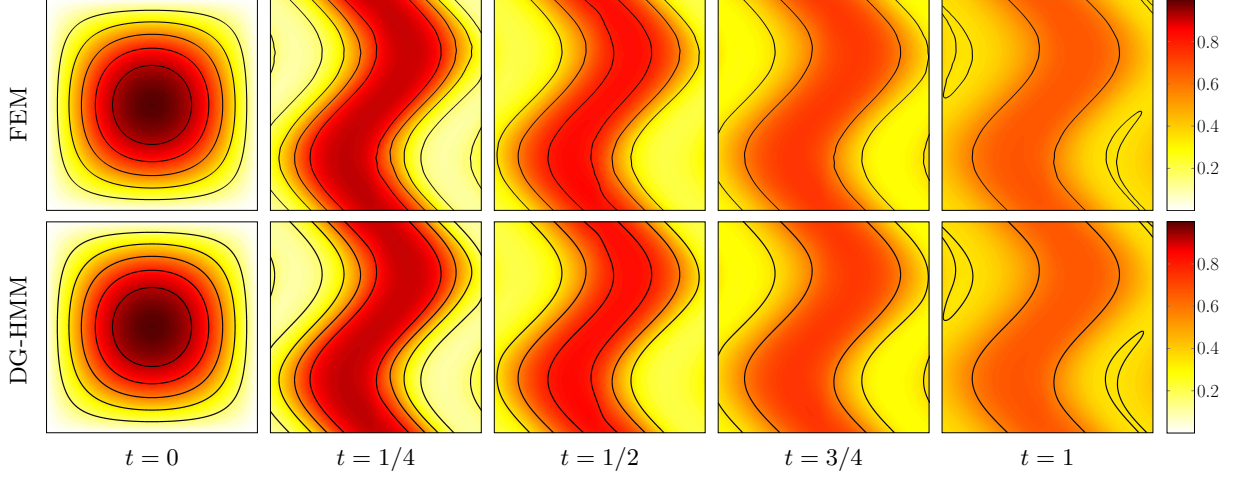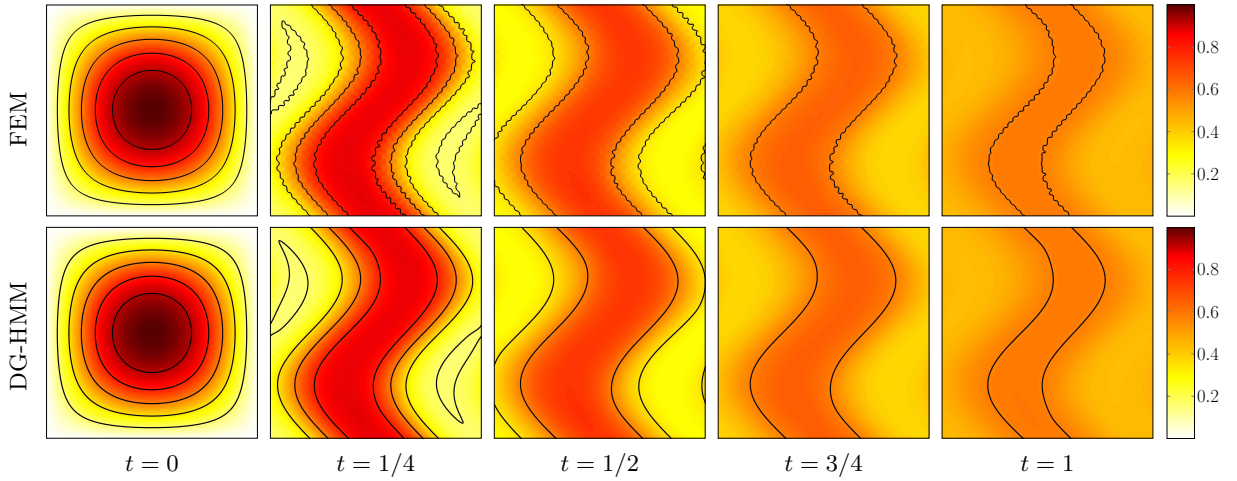
Figure 7: Test defined by (74) with $\varepsilon = 1/20$ and $\varsigma_{mac} = 1$, $\varsigma_{mic} = 0$, i.e., $b^\varepsilon(x) = b(x)$ without microscopic behavior. FE approximation of u^ε and DG-HMM solution u^H plotted on $\Omega = (0,1)^2$ at different times $t \in [0,T]$.

Figure 8: Test defined by (74) with $\varepsilon = 1/20$ and $\varsigma_{mac} = 1$, $\varsigma_{mic} = 1$, $\vartheta = 1$, i.e., $b^\varepsilon(x) = b(x, \frac{x}{\varepsilon})$ with $\mathrm{div}_y\, b(x,y) = 0$. FE approximation of u^ε and DG-HMM solution u^H plotted on $\Omega = (0,1)^2$ at different times $t \in [0,T]$.

Comparing Figure 4 with Figure 5, we observe that adding an incompressible periodic flow to a purely diffusive problem enhances the effective diffusion, while a compressible periodic flow, see Figure 6, yields a non-zero effective drift (although $\int_Y b(y)dy = 0$) and a depleted effective diffusion leading to internal layers in u^ε (FEM), which are removed by numerical upscaling in u^H (DG-HMM).

Next, the purely macroscopic drift in Figure 7 induces a channel, whose form nor location is altered by the incompressible periodic micro flow added in Figure 8, which solely increases the diffusion. In contrast, the compressible periodic micro flow added in Figure 9 transports the channel from the right to the left and yields checkerboard patterns in u^ε (FEM) due to a depleted diffusion. Although, there is a small phase shift between u^ε and u^H in Figure 9, the effective behavior of u^ε is well-captured by the HMM solution in all Figures 4–9, even for the locally periodic tests, for which homogenization is not known yet.

As homogenization for (1) with non-periodic data is not yet fully understood, we cannot expect that the HMM solution u^H accurately approximates the effective dynamics of u^ε as for instance the effective model (1) or the formulas for the effective data (9) might be inappropriate. In what follows, we present two test problems with locally periodic data (one with incompressible and one with compressible micro flow), where u^H does *not* entirely capture the effective behavior of u^ε. For further comparisons, we additionally calculate an approximation u^H_{am} using the arithmetic mean as "effective" data, i.e., we set $a^{\mathrm{eff},h}_K = \int_Y a(x_K, y)dy$ and $b^{\mathrm{eff},h}_K = \int_Y \varepsilon^{-1} b(x_K, y)dy$ in (21) and (19), respectively.
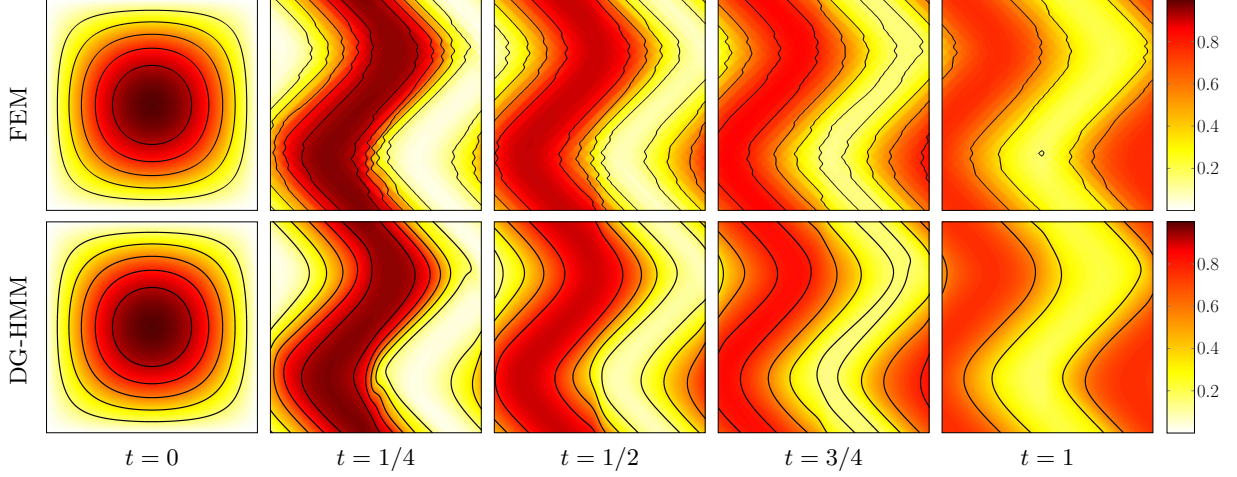
Figure 9: Test defined by (74) with $\varepsilon = 1/20$ and $\varsigma_{mac} = 1$, $\varsigma_{mic} = 1$, $\vartheta = -1$, i.e., $b^\varepsilon(x) = b(x, \frac{x}{\varepsilon})$ with $\text{div}_y\, b(x,y) \neq 0$. FE approximation of u^ε and DG-HMM solution u^H plotted on $\Omega = (0,1)^2$ at different times $t \in [0, T]$.

We first consider the data $a^\varepsilon(x) = a(x, x/\varepsilon)$, $b^\varepsilon(x) = \varepsilon^{-1}b(x, x/\varepsilon)$ with

$$a(x,y) = \frac{1}{1000}\left(\tfrac{11}{10} + \cos(2\pi x_1)\sin(2\pi y_1)\cos(2\pi y_2)\right)Id, \quad b(x,y) = \begin{pmatrix} \frac{1}{5}\sin^2(2\pi(x_1 + x_2 + y_2)) \\ 0 \end{pmatrix}. \quad (75)$$

Note that $\int_Y b(x,y)dy \neq 0$ and $\text{div}_y\, b(x,y) = 0$. In Figure 10, for $\varepsilon = 1/20$, we plot u^ε, u^H (obtained by (16)) and u_{am}^H (based on arithmetic means) for data (75). We observe that the diffusion at macro scale is adequately captured by u^H while the upscaled effective drift yields a phase shift. For u_{am}^H however, both macro diffusion and drift are wrong.

The result of Figure 10 might suggest, that mainly the definition of the numerically upscaled drift $b_K^{\text{eff},h}$ defined in (19) has to be adapted. However, for $a^\varepsilon(x) = a(x, x/\varepsilon)$, $b^\varepsilon(x) = \varepsilon^{-1}b(x, x/\varepsilon)$ with

$$\begin{aligned} a(x,y) &= \frac{1}{100}(3 + \sin(2\pi(x_1 + x_2)) + \sin(2\pi y_1)\cos(2\pi y_2))Id, \\ b(x,y) &= \frac{1}{4}\begin{pmatrix} \sin(2\pi x_1)\cos(2\pi x_2)\sin(2\pi y_1)\cos(2\pi y_2) \\ -\sin(2\pi x_1)\cos(2\pi y_1)\sin(2\pi y_2) \end{pmatrix}, \end{aligned} \quad (76)$$

(satisfying $\int_Y b(x,y)dy = 0$, $\text{div}_y\, b(x,y) \neq 0$), we also see in Figure 11, where u^ε, u^H (calculated by (16)) and u_{am}^H (based on arithmetic means) are plotted for $\varepsilon = 1/20$, that there are situations where the numerically upscaled diffusion is not enhanced enough (here, in particular, in the vertical direction). In contrast, we observe in u^ε that the effective drift seems to be zero everywhere, which is well captured by u^H. By comparison with u_{am}^H, that is completely wrong as no influence of b^ε on the macro diffusion is modeled, we again observe that our numerical homogenization method captures some essential features of the flow dynamics.

In summary, there are indeed situations where the multiscale method (16) does *not* reproduce all features of the macro dynamics of u^ε. However, the upscaling procedure used in (16) is clearly superior to naive upscaling using arithmetic means.

7   Conclusion and perspectives

In this article, we have considered parabolic advection-diffusion problems on a diffusive timescale with data rapidly varying at a small scale ε. We have introduced a space-discrete multiscale method which solves an effective equation (formulated in the original physical variables) using a discontinuous Galerkin finite element method in space. The a priori unknown effective data are recovered from micro simulations within cells of size comparable to ε. The computational cost for the space-discrete method is thus independent of the smallest scale in the problem. While the multiscale method is defined for multiscale problems without a particular structure (like periodicity or random stationarity) of the spatial heterogeneities, we gave an a priori error analysis of the numerical method for periodic data, which is the only
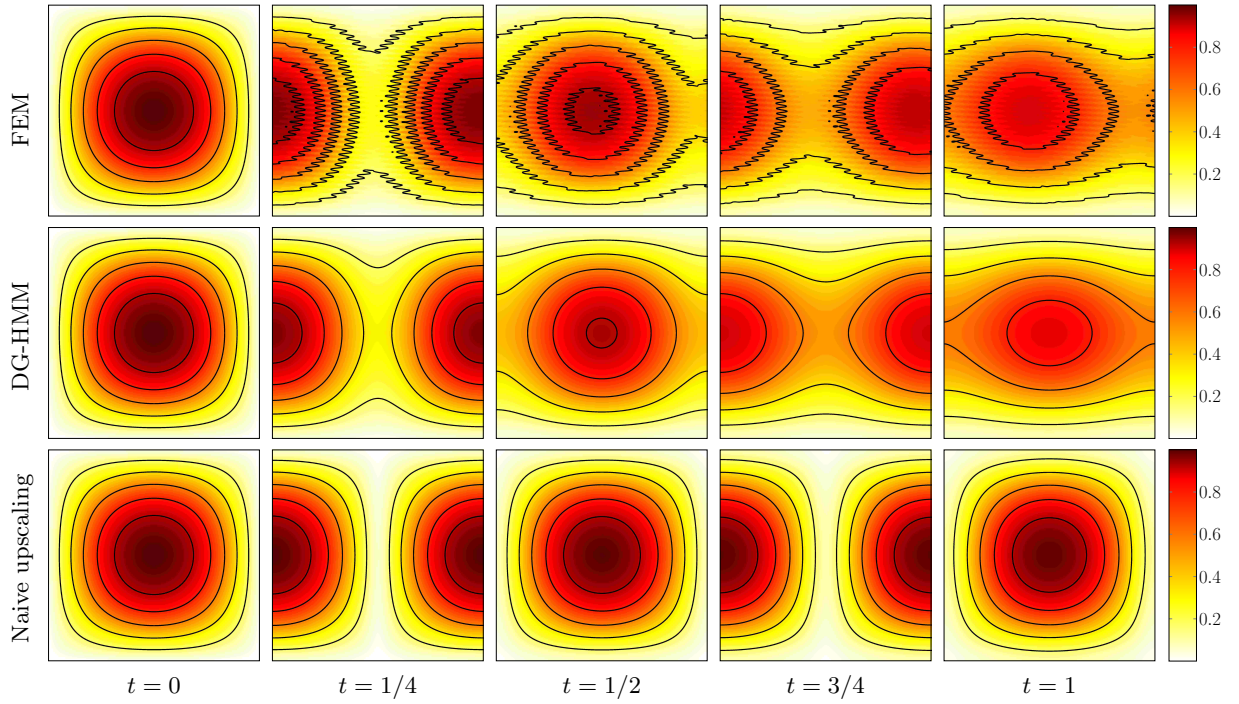
Figure 10: Test defined by (75) with $\varepsilon = 1/20$, i.e., $b^\varepsilon(x) = b(x, \frac{x}{\varepsilon})$ with $\text{div}_y\, b(x, y) = 0$. FE approximation of u^ε, DG-HMM solution u^H and naively upscaled solution u_{am}^H using arithmetic means plotted on $\Omega = (0, 1)^2$ at different times $t \in [0, T]$.
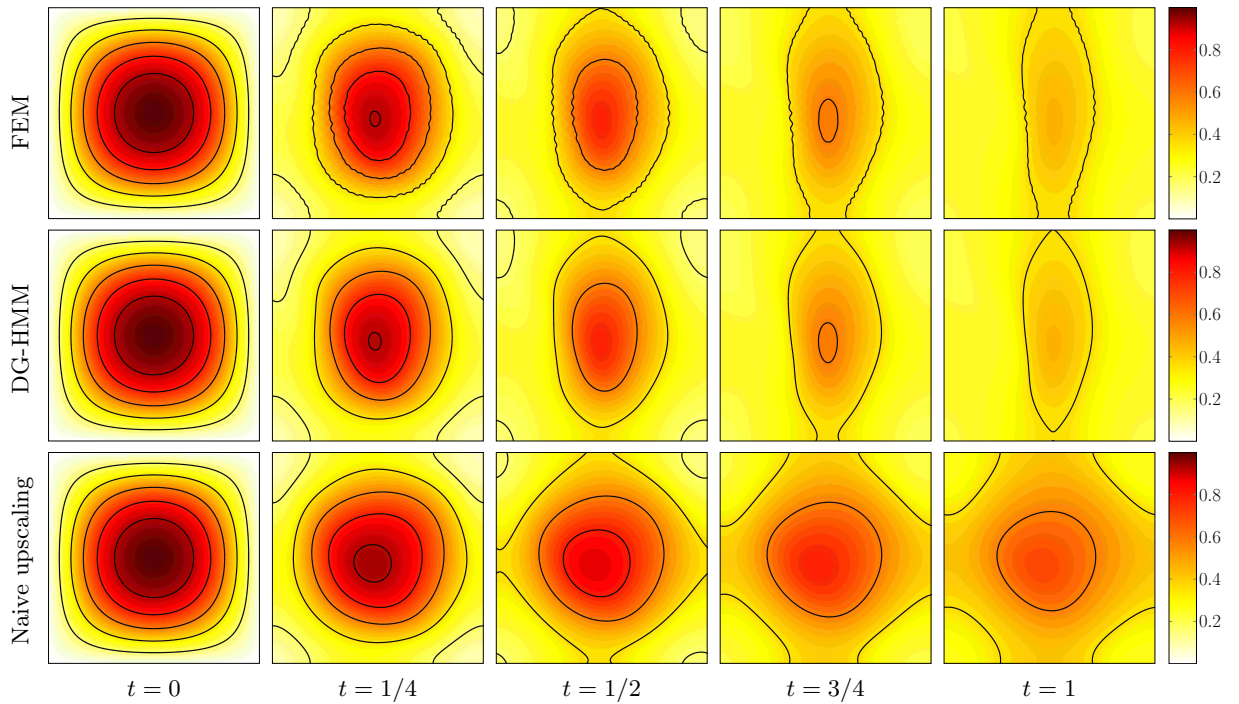

Figure 11: Test defined by (76) with $\varepsilon = 1/20$, i.e., $b^\varepsilon(x) = b(x, \frac{x}{\varepsilon})$ with $\text{div}_y\, b(x, y) \neq 0$. FE approximation of u^ε, DG-HMM solution u^H and naively upscaled solution u_{am}^H using arithmetic means plotted on $\Omega = (0, 1)^2$ at different times $t \in [0, T]$.

setting where homogenization results are available. The derived error estimates in the spatial DG norm are robust with respect to ε and give sharp convergence rates both for macro and micro spatial errors.

While in this article we focused on the space discretization and the upscaling strategy, an appropriate integrator in time is needed for practical implementation. Fully discrete space-time error estimates including time discretization error have yet to be derived. A computationally attractive choice would be the partitioned Runge-Kutta method PIROCK introduced in [9]. Further, extending our multiscale approach to more realistic transport models, e.g., allowing for inertial particles, to situations where the flow b^ε is determined by an associated Darcy problem, e.g., see [50, 54], or to transport models with additional reaction is of high interest for applied problems.

References

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

[2] ——, *The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs*, in Multiple scales problems in biomathematics, mechanics, physics and numerics, vol. 31 of GAKUTO Internat. Ser. Math. Sci. Appl., Gakkōtosho, Tokyo, 2009, pp. 133–181.

[3] ——, *A priori and a posteriori error analysis for numerical homogenization: a unified framework*, Ser. Contemp. Appl. Math. CAM, 16 (2011), pp. 280–305.

[4] ——, *Discontinuous Galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales*, Math. Comp., 81 (2012), pp. 687–713.

[5] A. ABDULLE AND S. ATTINGER, *Homogenization method for transport of DNA particles in heterogeneous arrays*, in Multiscale modelling and simulation, vol. 39 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2004, pp. 23–33.

[6] A. ABDULLE AND M. E. HUBER, *Finite element heterogeneous multiscale method for nonlinear monotone parabolic homogenization problems*. Submitted for publication. MATHICSE Technical Report 31.2014.

[7] ——, *Discontinuous Galerkin finite element heterogeneous multiscale method for advection–diffusion problems with multiple scales*, Numer. Math., 126 (2014), pp. 589–633.

[8] A. ABDULLE AND A. A. MEDOVIKOV, *Second order Chebyshev methods based on orthogonal polynomials*, Numer. Math., 90 (2001), pp. 1–18.

[9] A. ABDULLE AND G. VILMART, *PIROCK: a swiss-knife partitioned implicit-explicit orthogonal Runge-Kutta Chebyshev integrator for stiff diffusion-advection-reaction problems with or without noise*, J. Comput. Phys., (2013), pp. 869–888.

[10] ——, *Analysis of the finite element heterogeneous multiscale method for quasilinear elliptic homogenization problems*, Math. Comp., 83 (2014), pp. 513–536.

[11] S. AGMON, *Lectures on elliptic boundary value problems*, Van Nostrand Co, Princeton, N.J.-Toronto-London, 1965.

[12] G. ALLAIRE AND R. BRIZZI, *A multiscale finite element method for numerical homogenization*, Multiscale Model. Simul., 4 (2005), pp. 790–812 (electronic).

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

[14] G. ALLAIRE, I. PANKRATOVA, AND A. L. PIATNITSKI, *Homogenization and concentration for a diffusion equation with large convection in a bounded domain*, J. Funct. Anal., 262 (2012), pp. 300–330.

[15] 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), pp. 523–528.

[16] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

[17] B. Ayuso and L. D. Marini, *Discontinuous Galerkin methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1391–1420.

[18] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic analysis for periodic structures*, North-Holland Publishing Co., Amsterdam, 1978.

[19] F. Brezzi, L. D. Marini, and E. Süli, *Discontinuous Galerkin methods for first-order hyperbolic problems*, Math. Models Methods Appl. Sci., 14 (2004), pp. 1893–1903.

[20] E. Burman, *Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: Elliptic equations*, SIAM J. Sci. Comput., 35 (2013), pp. A2752–A2780.

[21] Y. Capdeboscq, *Homogenization of a diffusion equation with drift*, C. R. Acad. Sci. Paris Sér. I Math., 327 (1998), pp. 807–812.

[22] I. Capuzzo-Dolcetta and S. Finzi Vita, *Finite element approximation of some indefinite elliptic problems*, Calcolo, 25 (1988), pp. 379–395 (1989).

[23] C. Chainais-Hillairet and J. Droniou, *Finite-volume schemes for noncoercive elliptic problems with Neumann boundary conditions*, IMA J. Numer. Anal., 31 (2011), pp. 61–85.

[24] P. G. Ciarlet, *The finite element method for elliptic problems*, vol. 4 of Studies in Mathematics and its Applications, North-Holland, 1978.

[25] P. G. Ciarlet and P. A. Raviart, *Maximum principle and uniform convergence for the finite element method*, Comput. Methods Appl. Mech. Engrg., 2 (1973), pp. 17–31.

[26] P. Donato and A. L. Piatnitski, *Averaging of nonstationary parabolic operators with large lower order terms*, in Multi scale problems and asymptotic analysis, vol. 24 of GAKUTO Internat. Ser. Math. Sci. Appl., Gakkōtosho, Tokyo, 2006, pp. 153–165.

[27] R. Du and P. Ming, *Heterogeneous multiscale finite element method with novel numerical integration schemes*, Commun. Math. Sci., 8 (2010), pp. 863–885.

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

[29] L. C. Evans, *Partial Differential Equations*, Graduate studies in mathematics, American Mathematical Society, 1998.

[30] M. Feistauer, J. Hájek, and K. Švadlenka, *Space-time discontinuous Galerkin method for solving nonstationary convection-diffusion-reaction problems*, Appl. Math., 52 (2007), pp. 197–233.

[31] M. Feistauer and K. Švadlenka, *Discontinuous Galerkin method of lines for solving nonstationary singularly perturbed linear problems*, J. Numer. Math., 12 (2004), pp. 97–117.

[32] 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 09/09 - N, University of Münster, 2009.

[33] ——, *The heterogeneous multiscale finite element method for advection-diffusion problems with rapidly oscillating coefficients and large expected drift*, Netw. Heterog. Media, 5 (2010), pp. 711–744.

[34] T. R. Hill and W. H. Reed, *Triangular mesh methods for the neutron transport equation*, Los Alamos Report LA-UR-73-479, (1973).

[35] T. Y. Hou and P. J. Park, *Multiscale numerical methods for singularly perturbed convection-diffusion equations*, Int. J. Comput. Methods, 1 (2004), pp. 17–65.

[36] P. Houston, C. Schwab, and E. Süli, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 39 (2002), pp. 2133–2163.

[37] M. E. HUBER, *Numerical homogenization methods for advection-diffusion and nonlinear monotone problems with multiple scales*, PhD thesis, École Polytechnique Fédérale de Lausanne, Lausanne, 2015.

[38] V. V. JIKOV, *Asymptotic behavior and stabilization of solutions of a second-order parabolic equation with lowest terms*, Trudy Moskov. Mat. Obshch., 46 (1983), pp. 69–98.

[39] V. V. JIKOV, S. M. KOZLOV, AND O. A. OLEINIK, *Homogenization of differential operators and integral functionals*, Springer-Verlag, Berlin, Heidelberg, 1994.

[40] C. JOHNSON AND J. PITKÄRANTA, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comp., 46 (1986), pp. 1–26.

[41] J. KARÁTSON, S. KOROTOV, AND M. KŘÍŽEK, *On discrete maximum principles for nonlinear elliptic problems*, Math. Comput. Simulation, 76 (2007), pp. 99–108.

[42] K. KAVALIOU AND L. TOBISKA, *A finite element method for a noncoercive elliptic problem with Neumann boundary conditions*, Comput. Methods Appl. Math., 12 (2012), pp. 168–183.

[43] S. M. KOZLOV, *Reducibility of quasiperiodic differential operators and averaging*, Trudy Moskov. Mat. Obshch., 46 (1983), pp. 99–123.

[44] A. J. MAJDA AND P. R. KRAMER, *Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena*, Phys. Rep., 314 (1999), pp. 237–574.

[45] 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), pp. 391–409.

[46] G. NGUETSENG, *A general convergence result for a functional related to the theory of homogenization*, SIAM J. Math. Anal., 20 (1989), pp. 608–623.

[47] F. OUAKI, G. ALLAIRE, S. DESROZIERS, AND G. ENCHÉRY, *A priori error estimate of a multiscale finite element method for transport modeling*, SẽMA J., 67 (2015), pp. 1–37.

[48] F. OUAKI, G. ALLAIRE, G. ENCHÉRY, AND S. DESROZIERS, *A multiscale finite element method for transport modelling*, in CD-ROM Proceedings of the 6th European Congress on Computational Methods in Applied Sciences and Engineering, Vienna University of Technology, Austria, 2012.

[49] G. A. PAVLIOTIS, *Homogenization theory for advection-diffusion equations with mean flow*, PhD thesis, Rensselaer Polytechnic Institute, New York, 2002.

[50] G. A. PAVLIOTIS AND A. M. STUART, *Periodic homogenization for inertial particles*, Phys. D, 204 (2005), pp. 161–187.

[51] ———, *Multiscale Methods: Averaging and Homogenization*, vol. 53 of Text in Applied Mathematics, Springer-Verlag, New York, 2008.

[52] V. SOBOTÍKOVÁ, *Numerical integration in the DGFEM for 3D nonlinear convection-diffusion problems on nonconforming meshes*, Numer. Funct. Anal. Optim., 29 (2008), pp. 927–958.

[53] V. SOBOTÍKOVÁ AND M. FEISTAUER, *Effect of numerical integration in the DGFEM for nonlinear convection-diffusion problems*, Numer. Methods Partial Differential Equations, 23 (2007), pp. 1368–1395.

[54] S. SUN AND M. F. WHEELER, *Discontinuous Galerkin methods for coupled flow and reactive transport problems*, Appl. Numer. Math., 52 (2005), pp. 273–298.

[55] J. XU AND L. ZIKATANOV, *A monotone finite element scheme for convection-diffusion equations*, Math. Comp., 68 (1999), pp. 1429–1446.

[56] X. YUE AND W. E, *The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size*, J. Comput. Phys., 222 (2007), pp. 556–572.

[57] E. ZEIDLER, *Nonlinear functional analysis and its applications. II/A*, Springer-Verlag, New York, 1990. Linear monotone operators, Translated from the German by the author and Leo F. Boron.