Discontinuous Galerkin Finite Element Heterogeneous Multiscale Method for Advection-Diffusion Problems with Multiple Scales

A. Abdulle[1], M.E. Huber[2]

Abstract

A discontinuous Galerkin finite element heterogeneous multiscale method is proposed for advection-diffusion problems with highly oscillatory coefficients. The method is based on a coupling of a discontinuous Galerkin discretization for an effective advection-diffusion problem on a macroscopic mesh, whose a priori unknown data are recovered from micro finite element calculations on sampling domains within each macro element. The computational work involved is independent of the high oscillations in the problem at the smallest scale. The stability of our method (depending on both macro and micro mesh sizes) is established for both diffusion dominated and advection dominated regimes without any assumptions about the type of heterogeneities in the data. Fully discrete a priori error bounds are derived for locally periodic data. Numerical experiments confirm the theoretical error estimates.

Keywords: Heterogeneous multiscale method, discontinuous Galerkin methods, advection-diffusion, homogenization, variational crime, inf-sup condition

AMS subject classification (2010): 65N12, 65N15, 65N3, 74Q05.

1 Introduction

Transport problems from a wide range of applications including flow through aquifers in hydromecanic and ground water modeling, infiltration of contaminant in heterogeneous media, filtration processes and transport of chemicals in biological processes are modeled by advection-diffusion equations with highly heterogeneous coefficients varying on microscopic scales. However, the scale of interest in many applications is often much larger than these microscopic scales. Yet, standard numerical methods as the finite element method (FEM) are not able to capture the effective behavior of the transport phenomena unless the used discretization resolves the finest scale in the problem, denoted by ε in what follows. Such scale resolution is often prohibitive in terms of computational cost if the ratio between the finest scale and the scale of interest spans several orders of magnitude.

Analytic treatments of such problems often rely on the homogenization theory [19, 34] that aims at finding an averaged (effective) equation for the original heterogeneous equation in order to describe the effective behavior of the solution. General frameworks such as the G- or H-convergence (see [26] and [37], respectively) allow to homogenize partial differential equations (PDEs) without structure assumptions (e.g., such as periodicity or stationarity) on the heterogeneous coefficients. For the stationary advection-diffusion problems that we consider in this paper such results have been obtained in [13, 36]. We note that for non-stationary advection-diffusion problems (see for example [15]) the scaling for the advection term is usually different than the scaling used in [13, 36] for stationary problems. Indeed in [15] and related works, the model problem obtained from a diffusive scaling of an unscaled microscopic advection-diffusion equation, has a macroscopic Péclet number that scales as $\mathcal{O}(1/\varepsilon)$ (we recall that the Péclet number is defined as the ratio between the norms of the velocity field and the diffusion tensor of an advection-diffusion equation). In turn, the homogenized problem has an (advection) enhanced diffusion tensor. We note that the results are obtained for unbounded domains and in addition periodicity of the velocity and diffusion terms are assumed as well as a divergence free velocity. Such homogenization results cannot be applied to stationary advection-diffusion problems as considered in this paper. In contrast in the model [13, 36], the homogenized diffusion tensor will not depend on the advection term, a somehow less physical situation, but we allow for a non divergence free velocity and non-periodic velocity and diffusion terms.

[1]ANMC, Mathematics Section, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, assyr.abdulle@epfl.ch.

[2]ANMC, Mathematics Section, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, martin.huber@epfl.ch.

For practical purpose, the effective equations obtained by these analytic procedures are usually unknown and not available in explicit form and numerical methods are needed. For elliptic and parabolic homogenization problems, various multiscale approaches have been proposed – see [6, 9] and the references therein. In [2], using the methodology of the heterogeneous multiscale method (HMM) proposed by E and Engquist in [28], a time-dependent advection-diffusion problem where the diffusion tensor is a multiple of the unit tensor and the velocity field exhibits scale separation and high oscillations is studied. Therein, a heterogeneous multiscale method based on a standard finite element method is combined with a stabilized explicit Runge-Kutta method (ROCK) proposed in [10, 1]. Further, for one-dimensional hyperbolic and parabolic problems, a heterogeneous multiscale method based on a discontinuous Galerkin method has been proposed and analyzed in [21]. Moreover, a time-dependent advection-diffusion problem with large expected drift has been studied in [30] and in [32] numerical methods for singularly perturbed time-dependent advection-diffusion problems that exhibit a multiscale behavior have been considered. Finally, we also mention a numerical method based on the HMM that has been proposed in [40] for advection-diffusion problems. Similar stabilization techniques as used in this paper have been applied therein. However, the stability of the method has not been analyzed in [40]. This constitutes one of the main contribution of our paper. Furthermore, the a priori estimates derived there do not take into account the discretization error originating from the numerical computation of the effective data, i.e., the micro solver error.

In this paper we construct and analyze a numerical method for stationary advection-diffusion problems with highly oscillatory diffusion tensor and velocity field. The effective equation for our problem can be shown to be again an advection-diffusion problem with usually unknown effective diffusion tensor and velocity field [13, 36]. The Péclet number is allowed to be large, i.e., we consider the case of advection dominated problems. Two major modeling issues arise when trying to apply the methods developed for multiscale diffusion problems for (stationary) advection-diffusion problems. First, if the velocity field varies over multiple scales, proper upscaling for the advection term in the equation has to be built in the method. Second, even for constant velocity fields, when the methods developed for diffusion problems can be applied straightforwardly, the numerical method will become unstable if the Péclet number becomes large. This phenomenon is well-known for single scale problems and a large variety of stabilization techniques have been developed in order to adapt numerical methods to advection dominated problems (see [38] and references therein). One class of stabilized methods relies on discontinuous finite element methods (DG-FEM). In contrast to standard finite elements, the interelement continuity of the test and trial functions is relaxed, allowing flexibility in meshing and local conservation properties. First introduced in [31] for a purely hyperbolic transport equation, discontinuous Galerkin methods have been adapted to elliptic problems (see [17] for an overview) and, finally, to advection-diffusion-reaction problems. Their intrinsic flexibility makes them convenient for adaptivity techniques like local mesh refinement (h-refinement) or local increase of the polynomial order (p-refinement). An analysis of a hp-DG-FEM for advection-diffusion-reaction problems is presented in [33] by Houston, Schwab and Süli. Further, Ayuso and Marini proposed in [18] an analysis for a class of discontinuous Galerkin methods for advection-diffusion-reaction problems "relaxing" the usual coercivity condition relating the variable advection and reaction (see [18, Eq. (2.2)]) (as seen in [18], under these conditions, the analysis already for single scale problems is nontrivial). This is the setting adopted in this paper (see Section 5.1.1 for a discussion of this "relaxed coercivity condition").

The method proposed in this paper is built in the framework of the HMM. It is based on macroscopic discontinuous Galerkin (DG) FEM with input data (the effective diffusion tensor and the effective velocity field) recovered from micro FEM on sampling domains within the macroscopic elements. These sampling domains themselves scale with ε, the smallest scale in the problem and therefore, the computational work needed for our method is *independent* of the fine-scale features of the medium. Such combination of the HMM with the DG-FEM has already been proposed in [7] for pure diffusion problems. The extension to advection-diffusion problems is nontrivial. First, a careful (simultaneous) upscaling of the diffusion tensor and the velocity field has to be built in the method. Indeed, the case of an oscillatory velocity field, whose derivatives are unbounded with respect to ε is usually forbidden in a single scale analysis (see hypothesis (H2) in [18]). Second, the proof of the stability of the method is quite involved as the macro-micro strategy used here introduces variational crimes (due to numerical integration and numerical upscaling) leading to a non-consistent method and in particular to the loss of Galerkin orthogonality. Such variational crimes, also called non-consistent perturbations, are not made in the single scale method proposed in [18] (of course their method is non-conforming as it is also based on a discontinuous Galerkin method). For purely diffusive problems, as analyzed for the DG-HMM in [7], difficulties due to non-consistent perturbations can be avoided by using appropriate quadrature formula and the uniform ellipticity of the multiscale tensor. In contrast, the stability analysis including advection necessitates an upper bound for the non-

consistent perturbations committed in the advective part. Beside the non-consistency, the variational crimes in our DG-HMM introduce two further difficulties: in contrast to [18], the uniform boundedness on an infinite dimensional functional space of the diffusive part of our method in a DG norm is not available (due to numerical integration). To overcome this issue we need an appropriate interpolation result (see Section 5.1.1). Finally, all the data involved in the DG macro scheme come from micro simulations by FEM introducing yet another discretization error (due to the numerical upscaling procedure) that has to be taken into account in both the stability result and the a priori error estimates (see again Section 5.1.1).

The stability of the DG-HMM is established for a general class of diffusion tensors and velocity fields, that is, for general micro structures (without particular spatial structure such as periodicity or stationarity) and for advection or diffusion dominated problems. The method developed here is also suitable for advection-diffusion problems with boundary layers. Finally, optimal a priori error estimates are established for locally periodic data. We note that our a priori error estimates rely on new a priori error estimates for single scale DG-FEM based on numerical integration.

Our paper is organized as follows. In Section 2 we introduce the considered advection-diffusion problem. Then, we derive the multiscale method in Section 3. In Section 4, we summarize our main results on the stability and the a priori error estimates of the proposed method, the stability of the method for general data and the a priori estimates for locally periodic data. The proofs of the main results are provided in Section 5. Further, we present numerical experiments for both periodic and non-periodic (random) problems in Section 6 in order to confirm the theoretical estimates and to illustrate the capabilities of the proposed method. Finally, the results about the effect of numerical integration for single scale DG-FEM are derived in Appendix A.

Notation 1.1. In what follows, C denotes a generic positive constant, whose value can change at any occurrence (see also Remark 4.2). We consider the usual Sobolev spaces $W^{s,p}(\Omega)$. For $p = 2$, we use the notation $H^s(\Omega)$, $H^1_0(\Omega)$ for $p = 2$ and $s = 1$ with a vanishing trace on the boundary $\partial\Omega$, $W^1_{per}(Y) = \{v \in H^1_{per}(Y) \mid \int_Y v(y)dy = 0\}$ where $H^s_{per}(Y)$ is defined as the closure of $C^\infty_{per}(Y)$ (the subset of $C^\infty(\mathbb{R}^d)$ of periodic functions in $Y = (0,1)^d$) for the H^s norm, and $H^1(\mathrm{div}, \Omega) = \{v \in (L^2(\Omega))^d \mid \mathrm{div}\, v \in L^2(\Omega)\}$. For a matrix $a \in \mathbb{R}^{d\times d}$ with entries a_{ij}, we denote its Frobenius norm by $\|a\|_{\mathcal{F}} = \sqrt{\sum_i \sum_j |a_{ij}|^2}$. Further, we denote the vectors of the canonical basis of $\mathbb{R}^d$ by e_i, for $i = 1, \ldots, d$.

2 Model problem and homogenization results

In this section we introduce the advection-diffusion model problem and recall the results of classical homogenization theory.

Let $\Omega \subset \mathbb{R}^d$, $d \leq 3$, be a convex polygonal domain. Consider the multiscale advection-diffusion problem with zero Dirichlet boundary conditions

$$\begin{aligned} -\mathrm{div}(a^\varepsilon(x)\nabla u^\varepsilon(x)) + b^\varepsilon(x)\cdot\nabla u^\varepsilon(x) &= f(x) \qquad \text{in } \Omega, \\ u^\varepsilon(x) &= 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{1}$$

with $a^\varepsilon \in (L^\infty(\Omega))^{d\times d}$ the diffusion tensor, $b^\varepsilon \in (L^\infty(\Omega))^d$ the velocity field and the source $f \in L^2(\Omega)$. The variable ε represents a small scale in the problem, at which the data a^ε, b^ε fluctuates. We assume that the family of tensors a^ε (indexed by ε) are uniformly bounded and elliptic and the family of velocity fields b^ε is uniformly bounded, i.e., there exist $\lambda, \Lambda, \mathcal{B} > 0$ such that for any $\xi \in \mathbb{R}^d$

$$\lambda|\xi|^2 \leq a^\varepsilon(x)\xi\cdot\xi, \quad |a^\varepsilon(x)\xi| \leq \Lambda|\xi|, \quad |b^\varepsilon(x)| \leq \mathcal{B} \qquad \text{a.e. } x \in \Omega, \forall\, \varepsilon > 0. \tag{2}$$

As mentioned in the introduction, other scaling for the velocity field are usually used for non-stationary problems (i.e., velocity fields that are not uniformly bounded in ε, see e.g., [15]).

If the family of differential operators $-\mathrm{div}(a^\varepsilon(x)\nabla\cdot) + b^\varepsilon(x)\cdot\nabla\cdot$ satisfies the coercivity

$$\int_\Omega a^\varepsilon(x)\nabla v\cdot\nabla v + b^\varepsilon(x)\cdot\nabla v\, v\, dx \geq \lambda\|\nabla v\|^2_{L^2(\Omega)} \qquad \forall v \in H^1_0(\Omega), \forall\varepsilon > 0, \tag{3}$$

an application of the Lax-Milgram theorem provides a unique weak solution $u^\varepsilon \in H^1_0(\Omega)$ for all $\varepsilon > 0$. These solutions are uniformly bounded independently of ε by $\|u^\varepsilon\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}$ implying the existence of a subsequence of $\{u^\varepsilon\}$ weakly convergent in $H^1_0(\Omega)$. As a result of the H-convergence [13, 36] there exists

a subsequence of $\{a^\varepsilon, b^\varepsilon\}$ such that the corresponding sequence of solutions $\{u^\varepsilon\}$ weakly converges to u^0 in $H_0^1(\Omega)$, the solution of the homogenized problem

$$\begin{aligned} -\operatorname{div}(a^0(x)\nabla u^0(x)) + b^0(x)\cdot\nabla u^0(x) &= f(x) && \text{in } \Omega, \\ u^0(x) &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{4}$$

where $a^0(x) \in (L^\infty(\Omega))^{d\times d}$ is the homogenized tensor and $b^0(x) \in (L^\infty(\Omega))^d$ the homogenized velocity field. Furthermore, a^0 satisfies again the uniform ellipticity of condition (2) and the operator $-\operatorname{div}(a^0(x)\nabla\cdot) + b^0(x)\cdot\nabla\cdot$ is coercive (in the sense of (3)). Thus, the homogenized problem (4) has a unique solution.

We note that the coercivity condition (3) often is ensured by the condition $-\operatorname{div} b^\varepsilon(x) \geq 0$ for a.e. $x \in \Omega$ and $\varepsilon > 0$, assuming the regularity $b^\varepsilon \in H^1(\operatorname{div}, \Omega)$. However, a similar condition for the homogenized velocity field, i.e., $-\operatorname{div} b^0 \geq 0$, does not follow from [13, Theorem 2.1] directly, but will be assumed later on in (18), as it is the standard assumption to analyze numerical methods for advection-diffusion problems.

We notice, that we have chosen zero Dirichlet boundary conditions for the model problem (1) for simplicity. The multiscale method and its analysis derived in Section 3 and Section 4 can be generalized to non-zero Dirichlet, Neumann or mixed boundary conditions.

Péclet number. To distinguish the different regimes of problem (4) we define the Péclet number $\operatorname{Pe}(x)$ for the homogenized problem (4) locally by

$$\operatorname{Pe}(x) = \frac{|b^0(x)|L}{\|a^0(x)\|_{\mathcal{F}}} \qquad \forall\, x \in \Omega,$$

where $L = \operatorname{diam}\Omega$ is the characteristic length of the domain. The problem (4) is called diffusion dominated, if $\operatorname{Pe}(x) \lesssim 1$ in Ω, or advection dominated, if $\operatorname{Pe}(x) \gg 1$ in Ω. We assume, that $\operatorname{Pe}(x)$ has the same order for any $x \in \Omega$ and we will hence define and use the global Péclet number Pe in the following (see Section 4 for the definition).

3 Multiscale method

The goal is to derive a multiscale method that captures the effective solution of (1) at lower computational cost than solving (1) with standard numerical methods. We recall that for problem (1) scale resolution with a standard numerical method involves a computational cost of $\mathcal{O}(\varepsilon^{-d})$. The method we want to build will involve a computational cost independent of ε. In [7] a multiscale method has been proposed for a purely diffusive multiscale problem applying a discontinuous Galerkin method on the macro scale. As input, this method uses solely the data provided by the fine scale tensor a^ε. After having set up the framework of [7], we explain how to extend the method to advection-diffusion multiscale problems. The proposed method will be able to capture the effective solution of the multiscale advection-diffusion problem by coupling a discontinuous Galerkin method based on quadrature points on the macro scale with a standard FEM on (micro) sampling domains centered at these quadrature points recovering the effective data.

3.1 Preliminaries

First, we introduce the quantities needed for the macroscopic discretization.

Macro mesh. We consider a macroscopic family of partitions $\mathcal{T}_H$ of Ω, where $\mathcal{T}_H$ is a shape-regular macro mesh on Ω consisting of open, simplicial, not curved elements K satisfying $\bigcup_{K\in\mathcal{T}_H} \overline{K} = \overline{\Omega}$. Furthermore, we assume for simplicity that $\mathcal{T}_H$ is a conforming macro mesh, i.e., there are no hanging nodes. The element diameter H_K is defined by $H_K = \operatorname{diam} K$ for $K \in \mathcal{T}_H$ and the macro mesh size H is defined by $H = \max_{K\in\mathcal{T}_H} H_K$. Here $H \gg \varepsilon$ is allowed. By $\mathcal{E}$ we denote the set of all open $(d-1)$-dimensional interfaces of the elements of $\mathcal{T}_H$. We decompose $\mathcal{E}$ into interior interfaces $\mathcal{E}_{int}$ and interfaces on the boundary $\mathcal{E}_B$, i.e., $\mathcal{E} = \mathcal{E}_{int} \cup \mathcal{E}_B$. The interface diameter H_e is defined by $H_e = \operatorname{diam} e$ for $e \in \mathcal{E}$. Furthermore, we use the notation

$$\int_\Gamma \cdot = \sum_{e\in\mathcal{E}} \int_e \cdot\,.$$

Based on the macro mesh $\mathcal{T}_H$ we introduce the piecewise Sobolev space

$$H^1(\mathcal{T}_H) = \prod_{K\in\mathcal{T}_H} H^1(K) = \{v \in L^2(\Omega) \mid v|_K \in H^1(K), \forall\, K \in \mathcal{T}_H\}.$$

Remark 3.1. For simplicity of exposition, we assume that the macro mesh is conforming. However, one of the most important advantages of discontinuous Galerkin methods is the admissibility of irregular meshes which are useful for local refinement of the mesh. The extension of the presented results to this class of more general meshes is straightforward.

Macro finite element space. We define the piecewise linear finite element space for the discontinuous Galerkin method on the macro mesh $\mathcal{T}_H$ by

$$V^1(\Omega, \mathcal{T}_H) = \{u^H \in L^2(\Omega) \,|\, u^H|_K \in \mathcal{P}^1(K) \text{ and } u^H|_{K\cap\partial\Omega} = 0, \forall\, K \in \mathcal{T}_H\},$$

where $\mathcal{P}^1(K)$ denotes the space of linear polynomials on $K \in \mathcal{T}_H$. We note, that $V^1(\Omega, \mathcal{T}_H) \subset H^1(\mathcal{T}_H)$. In comparison to a standard finite element space the interelement continuity of the finite element functions is not postulated. Therefore, in the process of deriving the weak formulation of equation (1) the integration by parts can only be applied piecewisely on elements $K \in \mathcal{T}_H$.

Quantities on interfaces. Due to the lack of interelement continuity non-cancelling terms on the boundaries ∂K arise which we need to quantify. Consider $e \in \mathcal{E}$, $v \in H^1(\mathcal{T}_H)$ and $b \in \mathbb{R}^d$. First, we denote the outer normal vector on Ω and the outer normal vector on K by n and n_K, respectively, for any $K \in \mathcal{T}_H$. We notice, that v is generally two-valued on ∂K in the sense of traces, for any $K \in \mathcal{T}_H$. If e is an *interior edge* $e \in \mathcal{E}_{int}$ then there exist two triangles $K_1, K_2 \in \mathcal{T}_H$ such that $\overline{K_1} \cap \overline{K_2} = \overline{e}$, provided with outer normal vector n_1 and n_2 as well as interior traces v_1 and v_2, respectively. The normal jump and the average of v on e are defined by

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

If $b \cdot n_1 \neq 0$ we define

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

The edge e is then called inflow and outflow edge of K_i and K_o with respect to b, respectively. For $b \cdot n_1 = b \cdot n_2 = 0$ we choose a unique definition for K_i and K_o in order to facilitate further notation. As $n_1 = -n_2$ and $|n_1| = 1$, there exists a smallest $k^* \in \{1, \dots, d\}$ such that $k^* = \min\{1 \le k \le d \,|\, (n_1)_k \neq (n_2)_k\}$. We then define

$$\begin{cases} K_i = K_1 \\ K_o = K_2 \end{cases} \text{ if } (n_1)_{k^*} > 0, \qquad \text{or} \qquad \begin{cases} K_i = K_2 \\ K_o = K_1 \end{cases} \text{ if } (n_1)_{k^*} \le 0.$$

Further, we introduce the simplified notation $n_i = n_{K_i}$.

Analogously, we define these quantities for boundary edges. If e is a *boundary edge* $e \in \mathcal{E}_B$ then there exists a triangle $K_1 \in \mathcal{T}_H$ such that $\overline{e} = \overline{K_1} \cap \partial\Omega$, provided with an outer normal vector n_1 being n and the interior trace v_1. The normal jump and the average of v on e are defined by $[\![v]\!] = v_1 \cdot n$ and $\{v\} = v_1$, respectively. If $b \cdot n < 0$ then $K_i = K_1, n_i = n$, if $b \cdot n \ge 0$ then $K_o = K_1, n_i = -n$ and furthermore

$$v_i = \begin{cases} v_1 & b \cdot n < 0 \\ 0 & b \cdot n \ge 0 \end{cases}, \qquad v_o = \begin{cases} v_1 & b \cdot n > 0 \\ 0 & b \cdot n \le 0 \end{cases}.$$

Remark 3.2. Since the numerical and the analytic velocity field may lead to different inflow and outflow directions, we will use the following notation: the indices i and o (e.g., K_i) will be used for the numerically computed velocity field, constant on one interface (later on denoted by $\{\overline{b^e}\}$), while for the analytic velocity field we use the indices $i_0(x)$ and $o_0(x)$, e.g., $K_{i_0(x)}$ for $x \in e \in \mathcal{E}$.

3.2 DG-HMM for advection-diffusion problems

In this section, we introduce a multiscale method for advection-diffusion problems based on a discontinuous Galerkin method on the macroscopic scale.

Quadrature formula. The proposed method is based on a quadrature formula on the macro mesh. As we consider piecewise linear macro finite elements on simplicial meshes, we apply a one-point quadrature formula, i.e., for a continuous function $g : \Omega \to \mathbb{R}$,

$$\int_\Omega g(x)dx \approx \sum_{K\in\mathcal{T}_H} |K| g(x_K), \tag{5}$$

where x_K is located at the barycenter of $K \in \mathcal{T}_H$. We note, that the quadrature formula (5) is exact for piecewise linear functions g.

Sampling domains. For the upscaling procedure of the micro scale data, sampling domains around the quadrature points are defined. Let $\delta \geq \varepsilon$. For a macro element $K \in \mathcal{T}_H$ we consider the sampling domain K_δ around the barycenter x_K defined by $K_\delta = x_K + \delta I$, where $I = (-\frac{1}{2}, \frac{1}{2})^d$.

Macro bilinear form. Having set up the necessary framework we define the macro bilinear form for the multiscale advection-diffusion problem. We consider the macro bilinear form B on $V^1(\Omega, \mathcal{T}_H) \times V^1(\Omega, \mathcal{T}_H)$ defined as the sum $B = B_D + B_A$ of the diffusive part B_D and the advective part B_A given by

$$\begin{aligned} B_D(v^H, w^H) &= \sum_{K \in \mathcal{T}_H} \frac{|K|}{|K_\delta|} \int_{K_\delta} a^\varepsilon \nabla v_K^h \cdot \nabla w_K^h dx - \int_\Gamma \{\overline{a^\varepsilon \nabla v^h}\} \cdot [\![w^H]\!] ds + \int_\Gamma \mu [\![v^H]\!] \cdot [\![w^H]\!] ds, \\ B_A(v^H, w^H) &= \sum_{K \in \mathcal{T}_H} \frac{|K|}{|K_\delta|} \int_{K_\delta} b^\varepsilon \cdot \nabla v_K^h \, w^H(x_K) dx - \int_\Gamma \{\overline{b^\varepsilon}\} \cdot [\![v^H]\!] w_i^H ds, \end{aligned} \tag{6}$$

where v_K^h, w_K^h are appropriate micro functions solving (8) on the sampling domain K_δ and the average of multiscale fluxes $\{\overline{a^\varepsilon \nabla v^h}\}$ is given by (9). The penalty weighting function μ is given piecewisely by $\mu|_e = \left\|\{a_K^0\}\right\|_{\mathcal{F}} \alpha H_e^{-1}$, where a_K^0 is the numerically approximated homogenized tensor that will be defined in (15) and the penalization parameter $\alpha > 1$ is a positive parameter independent of the local mesh size and the data $a^\varepsilon(x)$. On an interface $e \in \mathcal{E}$, the average of the effective velocity field $\{\overline{b^\varepsilon}\}$ is defined in (11) and the trace w_i^H is taken from the inflow element K_i with respect to $\{\overline{b^\varepsilon}\}$.

Micro solver. Let $K \in \mathcal{T}_H$ be a macro triangle and K_δ its sampling domain. On this sampling domain we consider a simplicial micro mesh $\mathcal{T}_h$ and the micro finite element space $S^q(K_\delta, \mathcal{T}_h)$ defined by

$$S^q(K_\delta, \mathcal{T}_h) = \{z^h \in W(K_\delta) \mid z^h|_T \in \mathcal{P}^q(T), \forall T \in \mathcal{T}_h\}, \tag{7}$$

where $\mathcal{P}^q(T)$ denotes the space of polynomials on the element T of total degree at most q, with $q \in \mathbb{N}_{>0}$, and where the choice of $W(K_\delta)$ determines the boundary conditions used for computing the micro functions v_K^h. We consider two different spaces:

- periodic coupling: $W(K_\delta) = W^1_{per}(K_\delta) = \left\{v \in H^1_{per}(K_\delta) \mid \int_{K_\delta} v \, dx = 0\right\}$;
- Dirichlet coupling: $W(K_\delta) = H^1_0(K_\delta)$.

The micro problems are defined by: find v_K^h such that $(v_K^h - v^H) \in S^q(K_\delta, \mathcal{T}_h)$ and

$$\int_{K_\delta} a^\varepsilon(x) \nabla v_K^h \cdot \nabla z^h dx = 0 \quad \forall \, z^h \in S^q(K_\delta, \mathcal{T}_h). \tag{8}$$

Average of multiscale fluxes. Similarly as in the DG-HMM for diffusion problems, proper averages of fluxes on edges are crucial for the bilinear form (6). We recall such a construction first introduced in [5, 7].

For $e \in \mathcal{E}_{int}$, there exist $K_1, K_2 \in \mathcal{T}_H$ with corresponding sampling domains K_δ^1, K_δ^2 such that $\overline{e} = \overline{K_1} \cap \overline{K_2}$. The average of multiscale fluxes on e is defined by

$$\{\overline{a^\varepsilon \nabla v^h}\} = \frac{1}{2}\left(\frac{1}{|K_\delta^1|} \int_{K_\delta^1} a^\varepsilon \nabla v_{K_1}^h dx + \frac{1}{|K_\delta^2|} \int_{K_\delta^2} a^\varepsilon \nabla v_{K_2}^h dx\right). \tag{9}$$

Further, for $e \in \mathcal{E}_B$, there exists $K \in \mathcal{T}_H$ with corresponding sampling domain K_δ such that $\overline{e} = \overline{K} \cap \partial\Omega$. The average of multiscale fluxes on e is defined by

$$\{\overline{a^\varepsilon \nabla v^h}\} = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon \nabla v_K^h dx.$$

We notice, that for the averages $\{\overline{a^\varepsilon \nabla v^h}\}$ we omit the index K for the micro functions v_K^h as the neighboring elements for any $e \in \mathcal{E}$ are well-defined. Further, we emphasize that in contrast to the usual definition of the flux average in DG methods, existence of traces is not required for the average of multiscale fluxes. We also notice that the computation of the average of multiscale fluxes does not lead to an extra computational effort, as the quantities involved in the above integrals need to be computed anyway in the formulation of the method (6).

Average of effective velocity field. In order to define the quantity $\{\overline{b^\varepsilon}\}$ we introduce additional notations. Let $K \in \mathcal{T}_H$ be a macro triangle and $\varphi_{K,0}^H, \ldots \varphi_{K,d}^H$ the collection of linear nodal basis functions

on K. Furthermore, let F_K denote an affine C^1-diffeomorphism such that $F_K(\hat{K}) = K$, where $\hat{K}$ is the simplicial reference element in $\mathbb{R}^d$. The nodal basis $\hat{\varphi}_0^H, \ldots, \hat{\varphi}_d^H$ on $\hat{K}$ is defined by $\hat{\varphi}_0^H(\hat{x}) = 1 - \sum_{i=1}^d \hat{x}_i$ and $\hat{\varphi}_i^H(\hat{x}) = \hat{x}_i$, for $i = 1, \ldots, d$. We order the basis functions on K such that $\varphi_{K,i}^H(F_K(\hat{x})) = \hat{\varphi}_i^H(\hat{x})$ for $i = 0, \ldots, d$ and $\hat{x} \in \hat{K}$. Let D_K be the Jacobian matrix of F_K which is a constant $d \times d$ matrix. We introduce the matrix $Q_{\varphi_K^h}$ given by

$$Q_{\varphi_K^h} = D_K \begin{pmatrix} | & & | \\ \nabla\varphi_{K,1}^h & \cdots & \nabla\varphi_{K,d}^h \\ | & & | \end{pmatrix}^T, \tag{10}$$

where $\varphi_{K,i}^h$ solves the micro problem (8) constrained by the macro nodal basis function $\varphi_{K,i}^H$ for $i = 1, \ldots, d$.

For $e \in \mathcal{E}_{int}$, there exist $K_1, K_2 \in \mathcal{T}_H$ with corresponding sampling domains K_δ^1, K_δ^2 such that $\overline{e} = \overline{K_1} \cap \overline{K_2}$. The average of the effective velocity field on e is defined by

$$\{\overline{b^\varepsilon}\} = \frac{1}{2}\left(\frac{1}{|K_\delta^1|} \int_{K_\delta^1} Q_{\varphi_{K_1}^h} b^\varepsilon(x)dx + \frac{1}{|K_\delta^2|} \int_{K_\delta^2} Q_{\varphi_{K_2}^h} b^\varepsilon(x)dx \right). \tag{11}$$

For $e \in \mathcal{E}_B$, there exists $K \in \mathcal{T}_H$ with corresponding sampling domain K_δ such that $\overline{e} = \overline{K} \cap \partial\Omega$. The average of the effective velocity field on e is defined by

$$\{\overline{b^\varepsilon}\} = \frac{1}{|K_\delta|} \int_{K_\delta} Q_{\varphi_K^h} b^\varepsilon(x)dx.$$

As for the average of multiscale fluxes, the average of the effective velocity field uses quantities already computed for assembling the stiffness matrix of the diffusive part B_D. In order to assemble the terms of B_A on the edges the effective velocity field has to be stored for every $K \in \mathcal{T}_H$ similarly as for the multiscale fluxes appearing in B_D.

Macro solution. Our multiscale method for computing an effective solution of problem (1) reads as follows: find $u^H \in V^1(\Omega, \mathcal{T}_H)$ such that

$$B(u^H, v^H) = \int_\Omega f v^H dx \quad \forall\, v^H \in V^1(\Omega, \mathcal{T}_H). \tag{12}$$

Remark 3.3. The diffusive part B_D is the extension of the incomplete interior penalty Galerkin method (IIPG) of [25] to multiscale problems. It only differs in the lack of the symmetrizing term in comparison to the multiscale method proposed in [7]. The advective part B_A is a multiscale extension of the advective part of the method proposed in [33].

3.3 A useful reformulation of the DG-HMM

For the stability and the a priori error analysis of the FE-HMM (multiscale methods based on standard FEM) it turns out that it is convenient to define a numerically homogenized tensor a_K^0 (cf. e.g. [7, Section 5]). In that way, the diffusive form B_D can be reformulated as a standard DG-FEM based on numerical integration applied to a modified macro problem. An analogous reformulation of the advective form B_A will also be derived (this will allow to motivate the definition of $\{\overline{b^\varepsilon}\}$). We emphasize that this reformulation will only be used for the analysis but not for actual numerical computations.

To begin with, we consider a micro problem with modified right-hand side for $i = 1, \ldots, d$: find $\psi_K^{i,h} \in S^q(K_\delta, \mathcal{T}_h)$ such that

$$\int_{K_\delta} a^\varepsilon(x) \nabla\psi_K^{i,h} \cdot \nabla z^h \, dx = -\int_{K_\delta} a^\varepsilon(x) e_i \cdot \nabla z^h dx \quad \forall\, z^h \in S^q(K_\delta, \mathcal{T}_h). \tag{13}$$

We also consider the following similar problem for $i = 1, \ldots, d$: find $\psi_K^i \in W(K_\delta)$ such that

$$\int_{K_\delta} a^\varepsilon(x) \nabla\psi_K^i \cdot \nabla z \, dx = -\int_{K_\delta} a^\varepsilon(x) e_i \cdot \nabla z \, dx \quad \forall\, z \in W(K_\delta). \tag{14}$$

Based on the finite element micro functions $\psi_K^{i,h}$ and the exact micro function ψ_K^i, we define two tensors

$$a_K^0 = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon\Big(Id + J_{\psi_K^h(x)}^T\Big)dx, \qquad \overline{a}_K^0 = \frac{1}{|K_\delta|} \int_{K_\delta} a^\varepsilon\Big(Id + J_{\psi_K(x)}^T\Big)dx, \tag{15}$$

where $J_{\psi^h_K(x)}$ and $J_{\psi_K(x)}$ are $d \times d$ matrices with entries $\left(J_{\psi^h_K(x)}\right)_{il} = \frac{\partial \psi^{i,h}_K}{\partial x_l}$ and $\left(J_{\psi_K(x)}\right)_{il} = \frac{\partial \psi^i_K}{\partial x_l}$, respectively. We recall that the tensor a^0_K enters in the penalty weighting function μ for the macro bilinear form (6). It can be computed following [6, Theorem 18]. The estimation of this tensor does not constitute a computational overhead as the micro problems (13) need already to be solved to assemble B_D (indeed the solution of (8) can be obtained from the solutions of (13)).

Lemma 3.4 ([7, Lemma 5.4, Corollary 5.5]). *Let v^h_K, w^h_K be the solutions of the micro problem (8) such that $v^h_K - v^H \in S^q(K_\delta, \mathcal{T}_h)$ (resp. $w^h_K - w^H \in S^q(K_\delta, \mathcal{T}_h)$) with $S^q(K_\delta, \mathcal{T}_h) \subset W^1_{per}(K_\delta)$ (periodic coupling) or $S^q(K_\delta, \mathcal{T}_h) \subset H^1_0(K_\delta)$ (Dirichlet coupling). Then the following identities hold*

$$\frac{1}{|K_\delta|}\int_{K_\delta} a^\varepsilon(x)\nabla v^h_K \cdot \nabla w^h_K dx = a^0_K \nabla v^H(x_K)\cdot\nabla w^H(x_K), \qquad \frac{1}{|K_\delta|}\int_{K_\delta} a^\varepsilon(x)\nabla v^h_K dx = a^0_K \nabla v^H(x_K).$$

The average of the multiscale fluxes can be reformulated analogously (similar to [7, Lemma 5.6]).

Lemma 3.5. *Let $K_1, K_2 \in \mathcal{T}_H$ having a common interface e. Let $v^h_{K_1}$ and $v^h_{K_2}$ be the solution of (8) in K_1 and K_2, respectively, constrained by $v^H \in V^1(\Omega, \mathcal{T}_H)$ employing periodic or Dirichlet boundary conditions. Then*

$$\{\overline{a^\varepsilon \nabla v^h}\} = \{a^0_K \nabla v^H\}.$$

Hence, one can reformulate the diffusive part B_D of the method (cf. [7, Proposition 5.7]) by

$$B_D(v^H, w^H) = \sum_{K\in\mathcal{T}_H} |K| a^0_K \nabla v^H(x_K)\cdot\nabla w^H(x_K) - \int_\Gamma \{a^0_K \nabla v^H\}\cdot[\![w^H]\!]ds + \int_\Gamma \mu [\![v^H]\!]\cdot[\![w^H]\!]ds. \quad (16)$$

Next, we define the velocity field b^0_K and its counterpart $\overline{b}^0_K$ analogously to a^0_K and $\overline{a}^0_K$, respectively,

$$b^0_K = \frac{1}{|K_\delta|}\int_{K_\delta} (Id + J_{\psi^h_K}(x)) b^\varepsilon(x) dx, \qquad \overline{b}^0_K = \frac{1}{|K_\delta|}\int_{K_\delta} (Id + J_{\psi_K}(x)) b^\varepsilon(x) dx.$$

Following Lemma 3.4 we have that $\frac{1}{|K_\delta|}\int_{K_\delta} b^\varepsilon \cdot \nabla v^h_K dx = b^0_K \cdot \nabla v^H(x_K)$, and if we set $v^H = \varphi^H_{K,i}$, observing that $\nabla \varphi^H_{K,i} = (D^T_K)^{-1} e_i$, we obtain

$$e_i^T D_K^{-1} b^0_K = b^0_K \cdot \nabla v^H = \frac{1}{|K_\delta|}\int_{K_\delta} (\nabla \varphi^h_i)^T b^\varepsilon dx,$$

for $i = 1, \dots, d$, hence, we obtain $b^0_K = \frac{1}{|K_\delta|}\int_{K_\delta} Q_{\varphi^h_K} b^\varepsilon dx$, where $Q_{\varphi^h_K}$ is defined in (10). Thus, it follows that $\{\overline{b^\varepsilon}\} = \{b^0_K\}$. Taking into account that for $K \in \mathcal{T}_H$ and $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$ it holds

$$\int_K b^0_K \cdot \nabla v^H w^H dx = |K|\, b^0_K \cdot \nabla v^H(x_K) w^H(x_K),$$

the advective part B_A can then be reformulated as

$$B_A(v^H, w^H) = \sum_{K\in\mathcal{T}_H} \int_K b^0_K \cdot \nabla v^H w^H dx - \int_\Gamma \{b^0_K\}\cdot[\![v^H]\!] w^H_i ds. \quad (17)$$

For the analysis of the method carried out in Section 4 and 5.2 we always use the macro bilinear form in its reformulated version of (16) and (17).

4 Main results

In this section, we present our main results, namely the stability of the multiscale method (6) for general data a^ε, b^ε and a priori error estimates in both advection and diffusion dominated regimes for locally periodic data. We start by defining the norm used in our analysis.

Definition 4.1. *For $v^H \in V^1(\Omega, \mathcal{T}_H)$, let the norm $|||v^H||| = (|||v^H|||_D^2 + |||v^H|||_A^2)^{1/2}$ be defined by the following problem-dependent norms*

$$|||v^H|||_D^2 = a_\infty \left\|\nabla v^H\right\|_{L^2(\Omega)}^2 + a_\infty \left|v^H\right|_{*,D}^2, \qquad |||v^H|||_A^2 = b_\infty \left\|v^H\right\|_{L^2(\Omega)}^2 + \left|v^H\right|_{*,A}^2,$$

where $a_\infty = \operatorname{ess\,sup}_{x\in\Omega} \left\|a^0(x)\right\|_{\mathcal{F}}$, $b_\infty = \operatorname{ess\,sup}_{x\in\Omega} \left|b^0(x)\right|$ and $\left|v^H\right|_{*,R}^2 = \sum_{e\in\mathcal{E}} \left\|\gamma(R)^{1/2}[\![v^H]\!]\right\|_{L^2(e)}^2$, a weighted mesh-dependent semi-norm with

$$\gamma(R) = \begin{cases} \alpha H_e^{-1} & \text{for } R = D, \text{ (i.e., with respect to diffusion),} \\ \left|b^0 \cdot n\right| & \text{for } R = A, \text{ (i.e., with respect to advection).} \end{cases}$$

We note that the diffusion norm $|||\cdot|||_D$ is the same as the norm used in [7] scaled by $a_\infty^{1/2}$. Hence, for an advection dominated problem $|||\cdot|||_D$ is dominated by $|||\cdot|||_A$. As in our multiscale method variational crimes are committed, we need to introduce two quantities $r_{vc,A,\mathcal{T}_H}$ and $r_{vc,A,\mathcal{E}}$, that quantify the variational crimes in B_A, due to numerical integration on the macroscopic scale and the numerical upscaling procedure. We consider

$$r_{vc,A,\mathcal{T}_H} = \frac{1}{b_\infty} \sup_{\substack{K\in\mathcal{T}_H \\ x\in K}} \left|b_K^0 - b^0(x)\right|, \quad r_{vc,A,\mathcal{E}} = \frac{1}{b_\infty} \sup_{\substack{e\in\mathcal{E} \\ x\in e}} \left|\{b_K^0\} - b^0(x)\right|,$$

and we define $r_{vc,A} = r_{vc,A,\mathcal{T}_H} + r_{vc,A,\mathcal{E}}$. Further, the global Péclet number for the effective problem is given by $\mathrm{Pe} = \frac{b_\infty L}{a_\infty}$, where $L = \operatorname{diam}\Omega$.

Remark 4.2. In order to have an analysis for which the constants do not blow up in either the diffusion or the advection dominated regimes, the dependence of usually generic constants on a_∞, b_∞ as well as $\lambda, \Lambda, \mathcal{B}$ cannot be neglected. Hence, we use generic constants which only depend on terms of the type $\frac{\Lambda}{\lambda}$, $\frac{a_\infty}{\lambda}$, $\frac{b_\infty}{\mathcal{B}}$ and quantities independent of $a_\infty, b_\infty, H, h, \varepsilon, \delta$.

4.1 Stability results

The stability of the multiscale method (6) is proved for data a^ε, b^ε without any special spatial structure. We recall that for data a^ε, b^ε satisfying (2) and (3), it holds for the effective data that $b^0 \in (L^\infty(\Omega))^d$ and (see [13, 36])

$$a^0 \in (L^\infty(\Omega))^{d\times d}, \qquad a^0(x)\xi\cdot\xi \geq \lambda|\xi|^2, \quad \left|a^0(x)\xi\right| \leq \Lambda|\xi| \qquad \text{a.e. } x\in\Omega, \forall\, \xi\in\mathbb{R}^d.$$

For the stability analysis, we assume additionally that

$$b^0(x) \in W^{1,\infty}(\Omega), \quad -\operatorname{div} b^0(x) \geq 0 \qquad \text{a.e. } x\in\Omega, \tag{18}$$

$$b^0(x) \text{ has no closed curves }, \qquad b^0(x) \neq 0 \qquad \forall\, x\in\Omega. \tag{19}$$

The condition (18) is the standard assumption for the velocity field for formulating and analyzing DG-FEM for single scale advection-diffusion problems. Further, the condition (19) is used in [18] for the construction of a weighting function φ (see equation (34)) fundamental for the analysis of DG-FEM for advection-diffusion-reaction problems in the $|||\cdot|||$ norm.

To prove the stability of the numerical method (6), we derive the following inf-sup condition.

Theorem 4.3. *Assume* (2), (18) *and* (19). *Then, there exist $\alpha > 1, H_0 > 0, R_0 > 0$ such that, for*

$$H < H_0, \qquad r_{vc,A} < R_0, \tag{20}$$

the numerical method defined by (6) *fulfills the inf-sup condition*

$$\sup_{w^H\in V^1(\Omega,\mathcal{T}_H)} \frac{B(v^H, w^H)}{|||w^H|||} \geq \alpha_S |||v^H||| \qquad \forall\, v^H \in V^1(\Omega,\mathcal{T}_H), \tag{21}$$

where the stability constant α_S is independent of $H, h, \varepsilon, \delta$.

We emphasize that the constants α, H_0, R_0 are independent of $H, h, \varepsilon, \delta$. Further, the inf-sup condition (21) implies the existence and uniqueness of the solution of (12).

Theorem 4.4. *Assume* (2), (18) *and* (19). *Then the bilinear form* (6) *is uniformly bounded on* $V^1(\Omega,\mathcal{T}_H)\times V^1(\Omega,\mathcal{T}_H)$ *independently of* H,h,ε,δ. *Furthermore, under the additional conditions* (20) *of Theorem 4.3, the problem* (12) *has a unique solution* u^H *in* $V^1(\Omega,\mathcal{T}_H)$ *which satisfies*

$$|||u^H||| \leq \frac{1}{\alpha_S}\|f\|_{L^2(\Omega)}, \tag{22}$$

where α_S, *independent of* H,h,ε,δ, *is the stability constant of Theorem 4.3.*

4.2 A priori error estimates

As discussed in Section 1, the a priori error estimates are derived for locally periodic data a^ε, b^ε. The estimates rely on new results about the effect of numerical integration for DG-FEM applied to single scale problems, see Appendix A. We note that the assumption of local periodic data is only needed to estimate the modeling error, whereas the estimates of the macro and micro error given in Lemma A.1 and Lemma 5.13, respectively, are still valid for non-periodic data.

Assumption 4.5. *We assume local periodicity of* a^ε *and* b^ε *in the sense that there exist a tensor* $a(x,y)$ *and a velocity field* $b(x,y)$ *both* Y*-periodic in* y *such that* $a^\varepsilon(x)=a(x,\frac{x}{\varepsilon})$ *and* $b^\varepsilon(x)=b(x,\frac{x}{\varepsilon})$. *Furthermore, we postulate Lipschitz continuity of* a *and* b *with respect to the first variable*

$$a_{ij}(x,y), b_i(x,y) \in W^{1,\infty}(\overline{\Omega},L^\infty(Y)), \qquad 1\leq i,j\leq d.$$

In order to derive the a priori error estimates, we aim to decompose the total error into macro error e_{mac} and the quantity e_{HMM}

$$|||u^0-u^H||| \leq e_{mac}+e_{HMM}, \tag{23}$$

where u^0 is the solution of the homogenized problem (4) and u^H is the solution of (12). The explicit formulas for e_{mac} and e_{HMM} will be given in the proof of Theorem 4.6. While the macro error is due to the macroscopic solver $\tilde{B}_0$, a discontinuous Galerkin finite element method with numerical quadrature (introduced in Section 5.2.1) to solve the homogenized problem (4), the term e_{HMM} is the error solely caused by the upscaling procedure intrinsically encoded into B. By a Strang type result similar to [7, Lemma 5.10] we can trace e_{HMM} back to the quantities

$$r_{HMM,D}=\frac{1}{a_\infty}\sup_{K\in\mathcal{T}_H}\left\|a_K^0-a^0(x_K)\right\|_{\mathcal{F}}, \qquad r_{HMM,A}=\frac{1}{b_\infty}\sup_{K\in\mathcal{T}_H}\left|b_K^0-b^0(x_K)\right|. \tag{24}$$

Theorem 4.6. *Let* u^0 *and* u^H *be the solutions of* (4) *and* (12), *respectively. Assume* (2), (18), (19) *and the regularity* $u^0\in H^2(\Omega)$, $a^0\in(W^{1,\infty}(\Omega))^{d\times d}$, $b^0\in(W^{2,\infty}(\Omega))^d$. *Furthermore, assume that* H, $r_{vc,A}$ *and* α *satisfy the hypothesis of Theorem 4.3. Then, the following a priori error estimate holds*

$$|||u^0-u^H||| \leq C\Big(a_\infty^{1/2}H+b_\infty^{1/2}H^{3/2}+b_\infty^{1/2}\min\left\{\mathrm{Pe}^{1/2}H^2,H\right\}+a_\infty^{1/2}r_{HMM,D}+b_\infty^{1/2}r_{HMM,A}\Big)\left\|u^0\right\|_{H^2(\Omega)},$$

where C *is independent of* H,h,ε,δ.

The fully discrete error analysis relies on the decomposition of $r_{HMM,D}$ and $r_{HMM,A}$ into modeling error $r_{mod,\cdot}$ and micro error $r_{mic,\cdot}$ due to diffusion D and advection A, respectively,

$$r_{HMM,D}\leq \frac{1}{a_\infty}\sup_{K\in\mathcal{T}_H}\left\|a^0(x_K)-\overline{a}_K^0\right\|_{\mathcal{F}}+\frac{1}{a_\infty}\sup_{K\in\mathcal{T}_H}\left\|\overline{a}_K^0-a_K^0\right\|_{\mathcal{F}} =: r_{mod,D}+r_{mic,D}, \tag{25}$$

$$r_{HMM,A}\leq \frac{1}{b_\infty}\sup_{K\in\mathcal{T}_H}\left|b^0(x_K)-\overline{b}_K^0\right|+\frac{1}{b_\infty}\sup_{K\in\mathcal{T}_H}\left|\overline{b}_K^0-b_K^0\right| =: r_{mod,A}+r_{mic,A}. \tag{26}$$

In order to estimate the micro error we assume the following regularity of the exact micro functions ψ_K^i and the velocity field b^ε.

(H1) $\psi_K^i\in H^{q+1}(K_\delta)$ and $\left|\psi_K^i\right|_{H^{q+1}(K_\delta)}\leq C\varepsilon^{-q}\sqrt{|K_\delta|}$ for $K\in\mathcal{T}_H$, $i=1,\dots,d$.

(B1) $b^\varepsilon\in W^{1,\infty}(\Omega)$ and $|b_i^\varepsilon|_{W^{1,\infty}(\Omega)}\leq C\mathcal{B}\varepsilon^{-1}$ for $i=1,\dots,d$.

We refer to [7, Remark 5.1] for justification and discussion of (**H1**). Further, we observe that for smooth periodic velocity fields of the form $b^\varepsilon(x) = b(x/\varepsilon) = b(y)$, Y-periodic in y, we have by the chain rule $\partial_{x_i} b(x/\varepsilon) = \frac{1}{\varepsilon}\partial_{y_i} b(y)$. For the analysis of the modeling error we assume

(H2) a^ε and b^ε are collocated in the slow variable at x_K, i.e., $a^\varepsilon(x) = a(x_K, x/\varepsilon)$ and $b^\varepsilon(x) = b(x_K, x/\varepsilon)$ on any $K \in \mathcal{T}_H$, where x_K is the quadrature node in K.

This is possible due to the local periodicity of a^ε and b^ε (see Remark 5.16 for generalization).

For the discussion of the micro errors $r_{mic,\cdot}$ and the modeling errors $r_{mod,\cdot}$ we refer to Section 5.2.2. The fully discrete a priori error bounds for the DG-HMM proposed in this paper read as follow.

Theorem 4.7. *In addition to the assumptions of Theorem 4.6, assume (**H1**), (**H2**) and (**B1**). Then*

- *for periodic coupling, i.e., $W(K_\delta) = W^1_{per}(K_\delta)$ and $\frac{\delta}{\varepsilon} \in \mathbb{N}$,*

$$|||u^0 - u^H||| \leq C\left(a_\infty^{1/2} H + b_\infty^{1/2} H^{3/2} + b_\infty^{1/2} \min\left\{\mathrm{Pe}^{1/2} H^2, H\right\} + a_\infty^{1/2}\left(\frac{h}{\varepsilon}\right)^{2q} + b_\infty^{1/2}\left(\frac{h}{\varepsilon}\right)^{q+1}\right), \tag{27}$$

- *for Dirichlet coupling, i.e., $W(K_\delta) = H^1_0(K_\delta)$ and $\delta > \varepsilon$,*

$$\begin{aligned}|||u^0 - u^H||| \leq C\Big(a_\infty^{1/2} H + b_\infty^{1/2} H^{3/2} + b_\infty^{1/2} \min\left\{\mathrm{Pe}^{1/2} H^2, H\right\} + a_\infty^{1/2}\left(\frac{h}{\varepsilon}\right)^{2q} + a_\infty^{1/2}\frac{\varepsilon}{\delta} \\ + b_\infty^{1/2}\left(\frac{h}{\varepsilon}\right)^{q+1} + b_\infty^{1/2}\left(\frac{\varepsilon}{\delta}\right)^{1/2}\Big),\end{aligned} \tag{28}$$

where C is independent of $H, h, \varepsilon, \delta$.

We note that the first two terms of estimates (27) and (28) are known from the single scale analysis [18], the third term quantifies the effect of numerical integration on the macro scale derived in Theorem A.1 and the remaining terms describe the influence of micro and modeling error solely due to the multiscale strategy.

Remark 4.8. We observe that the numerical integration in the advective part introduces the additional term $b_\infty^{1/2} \min\left\{\mathrm{Pe}^{1/2} H^2, H\right\}$ into the a priori estimates (compare estimate (61) and Theorem A.1). We note that this additional term is at least of linear order. In our numerical experiments however, the linear order of convergence is never seen, not even for large Péclet number.

Further, the a priori estimate presented in Theorem 4.7 allows to define micro-macro refinement strategies for optimal convergence in the $|||\cdot|||$ norm with minimal computational costs for advection or diffusion dominated problems. Indeed, assume for example that we choose linear micro finite elements, i.e., $q = 1$, and we denote by N_{mac} and N_{mic} the number of macro and micro elements in each spatial dimension of the macro and the micro mesh, respectively, when discretizing Ω and the sampling domains K_δ by quasi-uniform triangular meshes (we note that the choice $q = 1$ leads to a quadratic convergence of $r_{mic,D}$ and $r_{mic,A}$ which is of higher order than the convergence of e_{mac}). We get the refinement strategies

(DD) $H \sim \left(\frac{h}{\varepsilon}\right)^2$ (i.e., $N_{mic} \sim \sqrt{N_{mac}}$) for diffusion dominated problems,

(AD) $H^{\frac{3}{2}} \sim \left(\frac{h}{\varepsilon}\right)^2$ (i.e., $N_{mic} \sim (N_{mac})^{3/4}$) for advection dominated problems.

Complexity. We recall, that N_{mac} and N_{mic} denote the number of macro and micro elements in each spatial dimension on the macro domain Ω and the micro domains K_δ, respectively. Thus, the macro and micro degree of freedom (DOF) are of order $\mathcal{O}(M_{mac})$ and $\mathcal{O}(M_{mic})$, respectively, where $M_{mac} = N_{mac}^d$ and $M_{mic} = N_{mic}^d$, and the total DOF is of order $\mathcal{O}(M_{mac} \cdot M_{mic})$. Further, the macro mesh size H is given by $H = 1/N_{mac}$ and the micro mesh size h by $h = \delta/N_{mic}$. Since the sampling domain size δ is of order $\mathcal{O}(\varepsilon)$ we get $h/\varepsilon = C/N_{mic}$, where C is a moderate constant. Therefore, in view of Theorem 4.7, N_{mic} can be chosen independently of ε.

The refinement strategies (DD) and (AD) presented above lead to optimal convergence rates in the $|||\cdot|||$ norm with minimal computational cost. We observe, that

$$M_{mic} \sim \sqrt{M_{mac}} \quad \text{for (DD) refinement,} \qquad M_{mic} \sim M_{mac}^{3/4} \quad \text{for (AD) refinement,}$$

leading to a complexity (independent of ε) of $\mathcal{O}(M_{mac}^{3/2})$ floating point operations for the optimal (linear) convergence rate in the $|||\cdot|||$ norm for a diffusion dominated problem and $\mathcal{O}(M_{mac}^{7/4})$ floating point operations for the optimal (superlinear) convergence rate in the $|||\cdot|||$ norm for an advection dominated problem. This holds under the assumption that the cost of the method is proportional to the total degree of freedom.

5 Proof of the main results

In this section, we provide the proofs of our main results presented in Section 4. We first prove the stability of the DG-HMM (6) in Section 5.1 and then derive the a priori error estimates in Section 5.2.

5.1 Proof of the stability results

In this section, we first show an interpolation result (Lemma 5.5) important for the stability proof and characterize the variational crimes in the advective part B_A of the DG-HMM (Lemma 5.7). Then, we give the proof of the stability result.

5.1.1 Bound on multiscale fluxes, weighting function, interpolation result and variational crimes

We start by stating some useful inequalities often used in what follows. The discrete Poincaré inequality (see [16, Lemma 2.1]) is given by

$$\|v\|_{L^2(\Omega)}^2 \leq C_P^2\Big(\|\nabla v\|_{L^2(\Omega)}^2 + |v|_{*,D}^2\Big) \qquad \forall\, v \in H^1(\mathcal{T}_H), \tag{29}$$

where $|\cdot|_{*,D}$ is the mesh-dependent semi-norm, introduced in Definition 4.1. Furthermore, we use two well-known results from standard finite element methods. The first one is the interpolation result of [22, Theorem 3.1.4]: let $K \in \mathcal{T}_H$, $k \in \mathbb{N}$, $u \in H^{k+1,p}(K)$ and $P_H u \in \mathcal{P}^k(K)$ be its L^2 orthogonal projection onto $\mathcal{P}^k(K)$. Then it holds

$$|u - P_H u|_{W^{m,q}(K)} \leq C|K|^{\frac{1}{q}-\frac{1}{p}} H_K^{k+1-m} |u|_{W^{k+1,p}(K)}, \tag{30}$$

for $1 \leq p, q \leq \infty$ and $m \in \mathbb{N}$ with $0 \leq m \leq k+1$. The second one is the scaled trace inequality [14, Theorem 3.10]: let $K \in \mathcal{T}_H$ and $e \in \mathcal{E}$ such that $e \subset \overline{K}$. Then, for $v \in H^1(K)$ we have

$$\|v\|_{L^2(e)} \leq C H_e^{1/2} \|\nabla v\|_{L^2(K)} + C H_e^{-1/2} \|v\|_{L^2(K)}. \tag{31}$$

In this article, we shall often use the combination of (31) and the inverse inequality [22, Theorem 3.2.6] for piecewise linear polynomials $v^H \in V^1(\Omega, \mathcal{T}_H)$

$$\big\|v^H\big\|_{L^2(e)} \leq C H_e^{-1/2} \big\|v^H\big\|_{L^2(K)}. \tag{32}$$

Finally, we recall that the quadrature formula (5) satisfies

$$\sqrt{\sum_{K\in\mathcal{T}_H} |K| |\nabla v^H(x_K)|^2} = \big\|\nabla v^H\big\|_{L^2(\Omega)}, \qquad \sqrt{\sum_{K\in\mathcal{T}_H} |K| |v^H(x_K)|^2} \leq C_{L2} \big\|v^H\big\|_{L^2(\Omega)},$$

for any $v^H \in V^1(\Omega, \mathcal{T}_H)$, where $C_{L2} > 0$ is a constant independent of H.

Bound on multiscale fluxes. The following energy inequalities and bounds on the multiscale fluxes related to the diffusive part B_D are needed for the stability proof. They have first been derived in [7] and [6]. First, we recall the energy equivalence [3, Proposition 3.2].

Lemma 5.1. *Let $v^H \in V^1(\Omega, \mathcal{T}_H)$ and let v_K^h be the solution of the micro problem* (8) *with either periodic or Dirichlet coupling. Assume that* (2) *holds, then*

$$\big\|\nabla v^H\big\|_{L^2(K_\delta)} \leq \big\|\nabla v_K^h\big\|_{L^2(K_\delta)} \leq \frac{\Lambda}{\lambda} \big\|\nabla v^H\big\|_{L^2(K_\delta)}.$$

As a simple consequence we recover uniform boundedness and ellipticity of the numerically homogenized tensor a_K^0 and a bound for the numerically homogenized velocity field b_K^0, respectively.

Corollary 5.2. *Let $v^H \in V^1(\Omega, \mathcal{T}_H)$, $K \in \mathcal{T}_H$ and assume (2). Then it holds*

$$|a_K^0 \nabla v^H(x_K)| \leq \frac{\Lambda^2}{\lambda} |\nabla v^H(x_K)|, \quad \|a_K^0\|_{\mathcal{F}} \leq \frac{\Lambda^2}{\lambda} d, \quad |b_K^0| \leq \mathcal{B}\frac{\Lambda}{\lambda}\sqrt{d}$$
$$a_K^0 \nabla v^H(x_K) \cdot \nabla v^H(x_K) \geq \lambda |\nabla v^H(x_K)|^2, \quad \|\{a_K^0\}\|_{\mathcal{F}} \geq \lambda.$$

We also recall the following bound for the multiscale fluxes first derived in [7, Lemma 4.3] that is fundamental for proving stability of the DG-HMM proposed in [7].

Lemma 5.3. *For $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$ and assuming (2) it holds*

$$\left| \int_\Gamma \{\overline{a^\varepsilon \nabla v^h}\} \cdot [\![w^H]\!] ds \right| \leq C_g \frac{\Lambda^2}{\lambda} \alpha^{-1/2} \|\nabla v^H\|_{L^2(\Omega)} |w^H|_{*,D},$$

where C_g is independent of $H, h, \varepsilon, \delta, \alpha$.

Weighting function. For the stability analysis of the advective part B_A, the use of a weighting function φ will be crucial. We motivate conditions (18), (19) and the use of φ by briefly reviewing the stability analysis for single scale problems with advection denoted by $b(x)$: the importance of considering a weighting function for the control of B_A was already noticed in [35], where problems with constant velocity fields b are studied. For variable velocity fields, the standard coercivity condition used in literature (see e.g. [20, 33] with reaction term set to zero) is given by

$$\exists\, c_0 > 0 \quad \text{such that} \quad -\operatorname{div} b(x) \geq c_0 \quad \forall\, x \in \Omega, \tag{33}$$

leading to coercivity results of the type $B_A(v^H, v^H) \geq c_0 \|v^H\|^2_{L^2(\Omega)} + |v^H|^2_{*,A}$. On one hand, the condition (33) excludes constant velocity fields b. On the other hand, $c_0 = 0$ would not allow to control $B_A(v^H, v^H)$ from below by the L^2 norm. In contrast, the weaker condition (19) that we assume in our analysis, used in [18, 24] in combination with (18), allows for stability results (through an inf-sup condition) for a broader class of velocity fields $b(x)$ provided a suitable use of a weighting function.

To construct such a weighting function, we follow the derivation given in [18] for single scale problems. The hypotheses (19) imply the existence of $\eta \in W^{2,\infty}(\Omega)$ such that $b^0(x) \cdot \nabla \eta(x) \geq 2b_\infty$ for any $x \in \Omega$ (see [18, Remark 2.1]). Using this function η, we introduce

$$\omega(x) = \exp(-\eta(x)), \qquad \varphi(x) = \omega(x) + \kappa, \tag{34}$$

where $\kappa > 0$ and φ is called *weighting function*. As $\eta \in W^{2,\infty}(\Omega)$ there exists a constant $C_\omega > 0$ such that

$$\frac{1}{C_\omega} \leq \omega(x) \leq C_\omega, \qquad |\nabla \omega(x)| \leq C_\omega, \qquad \|\omega\|_{W^{2,\infty}(\Omega)} \leq C_\omega \qquad \forall\, x \in \overline{\Omega}. \tag{35}$$

To bound φ from below and above we introduce $\underline{\varphi}^* = \frac{1}{C_\omega} + \kappa$ and $\overline{\varphi}^* = C_\omega + \kappa$, respectively. Finally, we observe that the weighting function φ satisfies

$$-b^0 \cdot \nabla \varphi(x) \geq 2b_\infty \frac{1}{C_\omega} \qquad \forall\, x \in \Omega. \tag{36}$$

Remark 5.4. For periodic data a^ε, b^ε the velocity field b^0 is constant. Therefore, conditions (19) reduce to $b^0 \neq 0$ and a possible choice for η is given by $\eta(x) = 2\frac{b^0}{b_\infty} \cdot x$. Since Ω is bounded, the bounds (35) can be specified explicitly using $e^{-2R} \leq \omega(x) \leq e^{2R}$, where $R > 0$ satisfies $\Omega \subset B_R(0)$.

Interpolation results. The analysis of the single scale methods in [18] uses the uniform boundedness of the diffusive part (see [18, Equation (4.9)]) of the form

$$B_D(v^H, w) \leq C_d |\!|\!| v^H |\!|\!|_D |\!|\!| w |\!|\!|_D \qquad \forall\, v^H \in V^1(\Omega, \mathcal{T}_H),\, w \in \tilde{V},$$

where $\tilde{V} \subset H^1(\mathcal{T}_H)$ is an infinite-dimensional subspace. However, since B_D is based on the quadrature formula (5), such a uniform boundedness of B_D does not hold in general. The following lemma, based on interpolation results, is used to overcome this difficulty.

Lemma 5.5. *Let $v^H \in V^1(\Omega, \mathcal{T}_H)$, φ defined in equation (34) and $P_H(\varphi v^H)$ be the L^2 projection of φv^H onto $V^1(\Omega, \mathcal{T}_H)$. Then, assuming (2), it holds*

$$\left| \sum_{K \in \mathcal{T}_H} |K| a_K^0 \nabla v^H(x_K) \cdot \nabla (\varphi v^H - P_H(\varphi v^H))(x_K) \right| \leq C \frac{\Lambda^2}{\lambda} C_\omega \left\| \nabla v^H \right\|_{L^2(\Omega)} \left\| v^H \right\|_{L^2(\Omega)},$$

where C is independent of H and κ.

Proof. We observe that $\varphi v^H - P_H(\varphi v^H) = \omega v^H - P_H(\omega v^H)$ is independent of κ. An application of the interpolation result (30) and using locally the inverse inequality [22, Theorem 3.2.6] (thus the constant C will depend on the shape regularity of $\mathcal{T}_H$) leads to

$$\begin{aligned} \left| \nabla (\omega v^H - P_H(\omega v^H))(x_K) \right| &\leq \left| \omega v^H - P_H(\omega v^H) \right|_{W^{1,\infty}(K)} \leq C |K|^{-1/2} H_K \left| \omega v^H \right|_{H^2(K)} \\ &\leq C |K|^{-1/2} H_K \|\omega\|_{W^{2,\infty}(K)} (\left\| v^H \right\|_{L^2(K)}^2 + \left\| \nabla v^H \right\|_{L^2(K)}^2)^{1/2} \\ &\leq C \|\omega\|_{W^{2,\infty}(K)} |K|^{-1/2} \left\| v^H \right\|_{L^2(K)}, \end{aligned}$$

which, combined with Corollary 5.2, concludes the proof

$$\begin{aligned} &\left| \sum_{K \in \mathcal{T}_H} |K| a_K^0 \nabla v^H(x_K) \cdot \nabla (\varphi v^H - P_H(\varphi v^H))(x_K) \right| \leq \frac{\Lambda^2}{\lambda} \sum_{K \in \mathcal{T}_H} |K| \left| \nabla v^H(x_K) \right| \left| \nabla (\omega v^H - P_H(\omega v^H))(x_K) \right| \\ &\leq C \frac{\Lambda^2}{\lambda} C_\omega \left\| \nabla v^H \right\|_{L^2(\Omega)} \left\| v^H \right\|_{L^2(\Omega)}. \end{aligned}$$

□

Furthermore, we recall the results of [18, Lemma 4.2] and extend them in a straightforward way to the norms introduced in Definition 4.1.

Lemma 5.6. *Let $\varphi \in W^{2,\infty}(\Omega)$ be the function defined in equation (34). For $v^H \in V^1(\Omega, \mathcal{T}_H)$ let $P_H(\varphi v^H)$ be the L^2 projection of φv^H onto $V^1(\Omega, \mathcal{T}_H)$. Then it holds*

$$\left\| \varphi v^H - P_H(\varphi v^H) \right\|_{L^2(\Omega)} \leq C C_\omega H \left\| v^H \right\|_{L^2(\Omega)}, \qquad \left| \varphi v^H - P_H(\varphi v^H) \right|_{H^1(\Omega)} \leq C C_\omega \left\| v^H \right\|_{L^2(\Omega)}, \tag{37}$$

$$\left(\sum_{K \in \mathcal{T}_H} \left\| \varphi v^H - P_H(\varphi v^H) \right\|_{L^2(\partial K)}^2 \right)^{1/2} \leq C C_\omega H^{1/2} \left\| v^H \right\|_{L^2(\Omega)}, \tag{38}$$

as well as

$$\left| P_H(\varphi v^H) - \varphi v^H \right|_{*,D} \leq C C_\omega \alpha^{1/2} \left\| v^H \right\|_{L^2(\Omega)}, \tag{39}$$

$$\left\| P_H(\varphi v^H) - \varphi v^H \right\| \leq C_I C_\omega \alpha^{1/2} \left\| v^H \right\|, \tag{40}$$

where the generic constants C, C_I are independent of $H, h, \varepsilon, \delta, \alpha, \kappa, \omega$.

Variational crimes. In order to prepare the analysis of the advective part B_A we derive an important bound for the variational crimes committed in B_A.

Lemma 5.7. *Let $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$ and $g \in \mathcal{C}^0(\overline{\Omega})$, then*

$$\begin{aligned} \left| \sum_{K \in \mathcal{T}_H} \int_K b_K^0 \cdot \nabla v^H \, w^H g \, dx - \int_\Omega b^0 \cdot \nabla v^H \, w^H g \, dx \right| &\leq \|g\|_{L^\infty(\Omega)} b_\infty r_{vc,A,\mathcal{T}_H} \left\| \nabla v^H \right\|_{L^2(\Omega)} \left\| w^H \right\|_{L^2(\Omega)}, \\ \left| \int_\Gamma \{b_K^0\} \cdot [\![v^H]\!] (w^H g)_i ds - \int_\Gamma b^0 \cdot [\![v^H]\!] (w^H g)_{i_0(s)} ds \right| &\leq C_r \|g\|_{L^\infty(\Omega)} \alpha^{-1/2} b_\infty r_{vc,A,\mathcal{E}} \left| v^H \right|_{*,D} \left\| w^H \right\|_{L^2(\Omega)}, \end{aligned}$$

where C_r is independent of $H, h, \delta, \varepsilon$ and α.

Proof. While the first estimate can be obtained in a straightforward way we split the second term into I_1 and I_2 with

$$I_1 = \int_\Gamma (\{b_K^0\} - b^0(s)) \cdot [\![v^H]\!] w_i^H g \, ds, \qquad I_2 = \int_\Gamma b^0(s) \cdot [\![v^H]\!] (w_i^H - w_{i_0(s)}^H) g \, ds,$$

where we used the continuity of g. The term I_1 can be estimated applying the scaled trace inequality (32)

$$\begin{aligned}|I_1| &\leq \|g\|_{L^\infty(\Omega)}\alpha^{-1/2}b_\infty r_{vc,A,\mathcal{E}}\big|v^H\big|_{*,D}\left(\sum_{e\in\mathcal{E}}H_e\big\|w_i^H\big\|^2_{L^2(e)}\right)^{1/2}\\ &\leq C\|g\|_{L^\infty(\Omega)}\alpha^{-1/2}b_\infty r_{vc,A,\mathcal{E}}\big|v^H\big|_{*,D}\big\|w^H\big\|_{L^2(\Omega)}.\end{aligned}$$

Further, we observe that for any $x\in e\in\mathcal{E}$ it holds

$$\text{either}\quad w_i^H(x)-w_{i_0(x)}^H(x)=0,\ (\text{if } n_{i_0(x)}=n_i)\quad\text{or}\quad \big|b^0(x)\cdot n_{i_0(x)}\big|\leq b_\infty r_{vc,A,\mathcal{E}},\ (\text{if } n_{i_0(x)}=-n_i),\quad(41)$$

using $b^0(x)\cdot n_{i_0(x)}\leq 0$ and $\{b_K^0\}\cdot n_{i_0(x)}\geq 0$ in the latter case. Combining (41) with the scaled trace inequality (32) leads to the estimate of I_2

$$\begin{aligned}|I_2| &\leq \|g\|_{L^\infty(\Omega)}\sum_{e\in\mathcal{E}}\int_{\{x\in e\,|\,n_{i_0(x)}\neq n_i\}}\big|b^0(s)\cdot n_{i_0(s)}\big|\big|[\![v^H]\!]\big|\Big|w_i^H-w_{i_0(s)}^H\Big|ds\\ &\leq C\|g\|_{L^\infty(\Omega)}\alpha^{-1/2}b_\infty r_{vc,A,\mathcal{E}}\big|v^H\big|_{*,D}\left(\sum_{e\in\mathcal{E}}H_e\Big|\big|w_i^H\big|+\Big|w_{i_0(x)}^H\Big|\Big|^2_{L^2(e)}\right)^{1/2}\\ &\leq C\|g\|_{L^\infty(\Omega)}\alpha^{-1/2}b_\infty r_{vc,A,\mathcal{E}}\big|v^H\big|_{*,D}\big\|w^H\big\|_{L^2(\Omega)}.\end{aligned}$$

□

5.1.2 Inf-sup condition and stability result

The inf-sup condition (21) in Theorem 4.3 relies on a lower bound of $B(v^H,\varphi v^H)$, where φ is the smooth weighting function defined in (34).

Lemma 5.8. *Let B_D and B_A be the bilinear forms of* (6) *and the weighting function φ be given by* (34)*. Under the conditions* (2)*,* (18) *and* (19)*, it holds*

$$B_D(v^H,\varphi v^H)\geq\frac{\lambda}{a_\infty}\left(\underline{\varphi}^*-C_\omega\frac{\Lambda^2}{\lambda^2}C_{L2}C_P-\frac{\overline{\varphi}^*}{2}C_g\frac{\Lambda^2}{\lambda^2}\alpha^{-1/2}\right)|\!|\!|v^H|\!|\!|_D^2,\tag{42}$$

$$B_A(v^H,\varphi v^H)\geq\frac{1}{2C_\omega}|\!|\!|v^H|\!|\!|_A^2-\overline{\varphi}^*\frac{\mathrm{Pe}^{1/2}}{L^{1/2}}(r_{vc,A,\mathcal{T}_H}+C_r\alpha^{-1/2}r_{vc,A,\mathcal{E}})|\!|\!|v^H|\!|\!|_D|\!|\!|v^H|\!|\!|_A,\tag{43}$$

$$|\!|\!|\varphi v^H|\!|\!|\leq C_\omega\sqrt{2}(C_P+\kappa)|\!|\!|v^H|\!|\!|.\tag{44}$$

Proof. First, the diffusive part is decomposed into three terms

$$\begin{aligned}B_D(v^H,\varphi v^H)=&\sum_{K\in\mathcal{T}_H}\frac{|K|}{|K_\delta|}\int_{K_\delta}a^\varepsilon\nabla v_K^h\cdot\nabla(\varphi v^H)_K^h dx-\int_\Gamma\{\overline{a^\varepsilon\nabla v^h}\}\cdot[\![\varphi v^H]\!]ds\\ &+\int_\Gamma\mu[\![v^H]\!]\cdot[\![\varphi v^H]\!]ds=I_1+I_2+I_3,\end{aligned}$$

denoting by $(\varphi v^H)_K^h$ the solution of the micro problem constrained by φv^H. Since $\varphi\in C^1(\Omega)$ the micro problem is well-defined. Using the reformulation (16), in view of the estimates of Corollary 5.2, we obtain

$$\begin{aligned}I_1=&\sum_{K\in\mathcal{T}_H}|K|a_K^0\nabla v^H(x_K)\cdot\nabla v^H(x_K)\varphi(x_K)+\sum_{K\in\mathcal{T}_H}|K|a_K^0\nabla v^H(x_K)\cdot\nabla\varphi(x_K)v^H(x_K)\\ \geq&\,\underline{\varphi}^*\lambda\big\|\nabla v^H\big\|^2_{L^2(\Omega)}-C_\omega\frac{\Lambda^2}{\lambda}C_{L2}\big\|\nabla v^H\big\|_{L^2(\Omega)}\big\|v^H\big\|_{L^2(\Omega)}\\ \geq&\,\underline{\varphi}^*\lambda\big\|\nabla v^H\big\|^2_{L^2(\Omega)}-C_\omega\frac{\Lambda^2}{\lambda}C_{L2}C_P\Big(\big\|\nabla v^H\big\|^2_{L^2(\Omega)}+\big|v^H\big|^2_{*,D}\Big),\end{aligned}$$

where the discrete Poincaré inequality (29) is used. Further, we observe that $[\![\varphi v^H]\!]=\varphi[\![v^H]\!]$ due to the regularity of φ. Applying Lemma 5.3 and Corollary 5.2 on I_2 and I_3, respectively, leads to

$$\begin{aligned}|I_2|=\left|\int_\Gamma\{\overline{a^\varepsilon\nabla v^h}\}\cdot[\![v^H]\!]\varphi\,ds\right|&\leq\overline{\varphi}^*C_g\frac{\Lambda^2}{\lambda}\alpha^{-1/2}\big\|\nabla v^H\big\|_{L^2(\Omega)}\big|v^H\big|_{*,D}\\ &\leq\frac{\overline{\varphi}^*}{2}C_g\frac{\Lambda^2}{\lambda}\alpha^{-1/2}\Big(\big\|\nabla v^H\big\|^2_{L^2(\Omega)}+\big|v^H\big|^2_{*,D}\Big),\\ I_3=\sum_{e\in\mathcal{E}}\big\|\{a_K^0\}\big\|_{\mathcal{F}}\int_e\alpha H_e^{-1}[\![v^H]\!]\cdot[\![v^H]\!]\varphi\,ds&\geq\underline{\varphi}^*\lambda\big|v^H\big|^2_{*,D}.\end{aligned}$$

For B_A, we first observe that the identity

$$\frac{1}{2}[\![(v^H)^2]\!] - [\![v^H]\!]v^H_{i_0(x)} = -\frac{1}{2}\big|[\![v^H]\!]\big|^2 n_{i_0(x)}, \tag{45}$$

holds for every $x \in e \in \mathcal{E}$. Then, we separate the terms due to variational crimes, apply integration by parts and use (45)

$$\begin{aligned}
B_A(v^H, \varphi v^H) = &\sum_{K\in\mathcal{T}_H}\int_K b^0\cdot\nabla v^H(\varphi v^H)dx - \int_\Gamma b^0\cdot[\![v^H]\!](\varphi v^H)_{i_0(s)}ds \\
&+ \underbrace{\sum_{K\in\mathcal{T}_H}\int_K (b^0_K - b^0)\cdot\nabla v^H(\varphi v^H)dx}_{I_4} - \underbrace{\int_\Gamma \{b^0_K\}\cdot[\![v^H]\!](\varphi v^H)_i - b^0\cdot[\![v^H]\!](\varphi v^H)_{i_0(s)}ds}_{I_5} \\
&= \frac{1}{2}\int_\Omega(-\operatorname{div}(b^0)\varphi - b^0\cdot\nabla\varphi)(v^H)^2dx + \int_\Gamma \varphi\, b^0\left(\frac{1}{2}[\![(v^H)^2]\!] - [\![v^H]\!]v^H_{i_0(s)}\right)ds + I_4 - I_5 \\
&\geq \frac{1}{C_\omega}b_\infty\int_\Omega (v^H)^2dx + \frac{1}{2}\int_\Gamma -b^0\cdot n_{i_0(s)}\big|[\![v^H]\!]\big|^2\varphi\, ds + I_4 - I_5,
\end{aligned}$$

where $-\operatorname{div} b^0 \geq 0$ of hypothesis (18) and the lower bound (36) are used. Using Lemma 5.7 with $g = \varphi$ for estimating the terms I_4 and I_5 leads to

$$\begin{aligned}
B_A(v^H, \varphi v^H) \geq &\frac{1}{C_\omega}b_\infty\big\|v^H\big\|^2_{L^2(\Omega)} + \frac{1}{2C_\omega}\big|v^H\big|^2_{*,A} \\
&- \overline{\varphi}^* b_\infty r_{vc,A,\mathcal{T}_H}\big\|\nabla v^H\big\|_{L^2(\Omega)}\big\|v^H\big\|_{L^2(\Omega)} - \overline{\varphi}^* C_r\alpha^{-1/2}r_{vc,A,\mathcal{E}}\big|v^H\big|_{*,D}\big\|v^H\big\|_{L^2(\Omega)},
\end{aligned}$$

from which the claimed lower bound (43) for the advective part B_A follows. Finally, the continuity of the mapping $v^H \mapsto \varphi v^H$ with respect to $|||\cdot|||$ can be shown by a direct computation. □

Since φv^H is not an element of $V^1(\Omega, \mathcal{T}_H)$ in general, we consider $w^H = P_H(\varphi v^H)$, where P_H denotes the L^2 projection onto $V^1(\Omega, \mathcal{T}_H)$. Hence, we decompose $B(v^H, P_H(\varphi v^H))$ into $B(v^H, \varphi v^H) + B(v^H, P_H(\varphi v^H) - \varphi v^H)$ where the second term can be seen as a perturbation.

Lemma 5.9. *Let B_D and B_A be as in* (6) *and φ be given by* (34)*. Under the conditions* (2), (18) *and* (19)*, there exist two positive constants C_D and C_A independent of ω and κ such that*

$$\begin{aligned}
\big|B_D(v^H, \varphi v^H - P_H(\varphi v^H))\big| &\leq C_D C_\omega\alpha^{1/2}|||v^H|||^2_D, \\
\big|B_A(v^H, \varphi v^H - P_H(\varphi v^H))\big| &\leq C_A C_\omega\alpha^{-1/2}H\frac{\mathrm{Pe}^{1/2}}{L^{1/2}}|||v^H|||_D|||v^H|||_A,
\end{aligned}$$

for any $v^H \in V^1(\Omega, \mathcal{T}_H)$, where C_D and C_A are independent of $H, h, \varepsilon, \delta, \alpha, \kappa$.

Proof. We start by estimating the diffusive part B_D

$$\begin{aligned}
B_D(v^H, P_H(\varphi v^H) - \varphi v^H) = &\sum_{K\in\mathcal{T}_H}|K|a^0_K\nabla v^H(x_K)\cdot\nabla(P_H(\varphi v^H) - \varphi v^H)(x_K) \\
&- \int_\Gamma\{\overline{a^\varepsilon\nabla v^h}\}\cdot[\![P_H(\varphi v^H) - \varphi v^H]\!]ds + \sum_{e\in\mathcal{E}}\big\|\{a^0_K\}\big\|_{\mathcal{F}}\int_e \alpha H_e^{-1}[\![v^H]\!]\cdot[\![P_H(\varphi v^H) - \varphi v^H]\!]ds.
\end{aligned}$$

We use Lemma 5.5, the interpolation estimate (39), Corollary 5.2 and $\alpha > 1$ to obtain the first estimate

$$\begin{aligned}
\big|B_D(v^H,\ P_H(\varphi v^H) - \varphi v^H)\big| &\leq C\frac{\Lambda^2}{\lambda}C_\omega\big\|\nabla v^H\big\|_{L^2(\Omega)}\big\|v^H\big\|_{L^2(\Omega)} \\
&\quad + C_g\frac{\Lambda^2}{\lambda}\alpha^{-1/2}\big\|\nabla v^H\big\|_{L^2(\Omega)}\big|P_H(\varphi v^H) - \varphi v^H\big|_{*,D} + \frac{\Lambda^2}{\lambda}d\big|v^H\big|_{*,D}\big|P_H(\varphi v^H) - \varphi v^H\big|_{*,D} \\
&\leq CC_\omega\frac{\Lambda^2}{\lambda}\Big((1+C_g)\big\|\nabla v^H\big\|_{L^2(\Omega)}\big\|v^H\big\|_{L^2(\Omega)} + \alpha^{1/2}\big|v^H\big|_{*,D}\big\|v^H\big\|_{L^2(\Omega)}\Big) \\
&\leq C_D C_\omega\alpha^{1/2}|||v^H|||^2_D.
\end{aligned}$$

For the advective part, we observe that the first term of $B_A(v^H, P_H(\varphi v^H) - \varphi v^H)$ (see (6)) vanishes

$$\int_K b_K^0 \cdot \nabla v^H (P_H(\varphi v^H) - \varphi v^H) dx = 0 \qquad \forall\, K \in \mathcal{T}_H,$$

using the definition of the L^2 projection and the fact that $b_K^0 \cdot \nabla v^H$ is constant on any $K \in \mathcal{T}_H$. Hence, applying Corollary 5.2 and the interpolation estimate (38) lead to

$$\begin{aligned}
\big|B_A(v^H, P_H(\varphi v^H) - \varphi v^H)\big| &= \int_\Gamma \big|\{b_K^0\}\big|\big|[\![v^H]\!]\big|\big|(P_H(\varphi v^H) - \varphi v^H)_i\big| ds \\
&\le C\mathcal{B}\alpha^{-1/2} H^{1/2} \big|v^H\big|_{*,D} \left(\sum_{e \in \mathcal{E}} \big\|(P_H(\varphi v^H) - \varphi v^H)_i\big\|^2_{L^2(e)} \right)^{1/2} \\
&\le C C_\omega \mathcal{B}\alpha^{-1/2} H \big|v^H\big|_{*,D} \big\|v^H\big\|_{L^2(\Omega)}. \qquad \square
\end{aligned}$$

Proof of Theorem 4.3. Let $v^H \in V^1(\Omega, \mathcal{T}_H)$. We consider $P_H(\varphi v^H) \in V^1(\Omega, \mathcal{T}_H)$, where the weighting function φ is defined in (34) up to the parameter $\kappa > 0$. For the proof of the inf-sup condition we prove

$$\text{(i)} \quad B(v^H, P_H(\varphi v^H)) \ge C_2 |\!|\!| v^H |\!|\!|^2, \qquad \text{(ii)} \quad |\!|\!| P_H(\varphi v^H) |\!|\!| \le C_1 |\!|\!| v^H |\!|\!|, \tag{46}$$

where $C_1, C_2 > 0$ are independent of H, h, ε and δ. These estimates then directly imply the inf-sup condition with stability constant $\alpha_S = C_2/C_1$

$$\sup_{w^H \in V^1(\Omega, \mathcal{T}_H)} \frac{B(v^H, w^H)}{|\!|\!| w^H |\!|\!|} \ge \frac{B(v^H, P_H(\varphi v^H))}{|\!|\!| P_H(\varphi v^H) |\!|\!|} \ge \frac{C_2 |\!|\!| v^H |\!|\!|^2}{C_1 |\!|\!| v^H |\!|\!|} = \alpha_S |\!|\!| v^H |\!|\!|.$$

For showing (46).(i), we combine Lemma 5.8, Lemma 5.9 and Young's inequality

$$\begin{aligned}
B_D(v^H, P_H(\varphi v^H)) &\ge \frac{\lambda}{a_\infty} \left(\underline{\varphi}^* - C_\omega \frac{\Lambda^2}{\lambda^2} C_{L2} C_P - \frac{\overline{\varphi}^*}{2} C_g \frac{\Lambda^2}{\lambda^2} \alpha^{-1/2} \right) |\!|\!| v^H |\!|\!|^2_D - C_D C_\omega \alpha^{1/2} |\!|\!| v^H |\!|\!|^2_D, \\
B_A(v^H, P_H(\varphi v^H)) &\ge \frac{1}{2C_\omega} |\!|\!| v^H |\!|\!|^2_A - \frac{1}{2} \frac{\mathrm{Pe}^{1/2}}{L^{1/2}} \Big(\overline{\varphi}^* (r_{vc,A,\mathcal{T}_H} + C_r \alpha^{-1/2} r_{vc,A,\mathcal{E}}) + C_A C_\omega \alpha^{-1/2} H \Big) |\!|\!| v^H |\!|\!|^2,
\end{aligned}$$

where the estimates are explicit in H, κ, α and $r_{vc,A,\mathcal{T}_H}, r_{vc,A,\mathcal{E}}$. The goal is to find constants $\alpha > 1$, $H_0 > 0$, $R_0 > 0$ and suitable choice of $\kappa > 0$, such that for $H < H_0$ and $r_{vc,A} < R_0$, it holds

$$\text{(a)} \quad B_D(v^H, P_H(\varphi v^H)) \ge 2C_* |\!|\!| v^H |\!|\!|^2_D, \qquad \text{(b)} \quad B_A(v^H, P_H(\varphi v^H)) \ge 2C_* |\!|\!| v^H |\!|\!|^2_A - C_* |\!|\!| v^H |\!|\!|^2. \tag{47}$$

We then immediately see that by setting $C_2 = C_*$ the estimate (46).(i) follows.

We thus start by proving (47).(a). First, we choose the penalization parameter α such that

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

is satisfied. Further, we choose $\kappa > 0$ such that the conditions

$$\mathtt{B}(\alpha, \kappa): \qquad \underline{\varphi}^* > \frac{\overline{\varphi}^*}{2}, \qquad \frac{\underline{\varphi}^*}{2} > C_\omega \frac{\Lambda^2}{\lambda^2} C_{L2} C_P, \qquad \frac{\underline{\varphi}^*}{8} \frac{\lambda}{a_\infty} > C_D C_\omega \alpha^{1/2},$$

hold, fixing the weighting function φ. Defining $C_* = \min\{\frac{C_D C_\omega \alpha^{1/2}}{2}, \frac{1}{4C_\omega}\}$ and for α and κ satisfying $\mathtt{A}$ and $\mathtt{B}$ the bound (47).(a) can be shown.

We continue with the conditions leading to the bound (47).(b). First, we define H_0 as the supremum over all $H > 0$ such that

$$\mathtt{C}(\alpha, H): \qquad \frac{1}{2} C_* > \frac{1}{2} \frac{\mathrm{Pe}^{1/2}}{L^{1/2}} C_A C_\omega H \alpha^{-1/2},$$

is satisfied. We thus have

$$B_A(v^H, P_H(\varphi v^H)) \ge 2C_* |\!|\!| v^H |\!|\!|^2_A - \left(\frac{1}{2} \frac{\mathrm{Pe}^{1/2}}{L^{1/2}} \overline{\varphi}^* \max(1, C_r) r_{vc,A} + \frac{1}{2} C_* \right) |\!|\!| v^H |\!|\!|^2.$$

Second, we define R_0 the value such that for $r_{vc,A} < R_0$ it holds

$$\mathtt{D}(\kappa, r_{vc,A}): \qquad \frac{1}{2}C_* > \frac{1}{2}\frac{\mathrm{Pe}^{1/2}}{L^{1/2}}\overline{\varphi}^* \max(1, C_r) r_{vc,A},$$

thus $B_A(v^H, P_H(\varphi v^H)) \geq 2C_* |\!|\!| v^H |\!|\!|_A^2 - C_* |\!|\!| v^H |\!|\!|^2$ for $H < H_0$ and $r_{vc,A} < R_0$ and we have shown (47).(b).

The second part (46).(ii) follows by using estimate (44) and the interpolation result (40)

$$|\!|\!| P_H(\varphi v^H) |\!|\!| \leq |\!|\!| \varphi v^H |\!|\!| + |\!|\!| P_H(\varphi v^H) - \varphi v^H |\!|\!| \leq C_1 |\!|\!| v^H |\!|\!|,$$

with $C_1 = C_\omega \max\{\sqrt{2}(C_P + \kappa), C_I \alpha^{1/2}\}$. □

We remark that any $\alpha > \alpha_0 > 1$, α_0 being a threshold value satisfying A, can be chosen. However, the choice of α influences κ through the condition B and R_0 through condition D. Hence, the choice of α has an impact on the upper bounds for the variational crimes $r_{vc,A}$.

Remark 5.10. It might be of interest to compare the conditions A, B, C, D to the conditions used in the stability proof of DG-FEM for single scale problems presented in [18, Theorem 4.4]. While the condition A corresponds to [18, Equation (3.2), (4.16)], B corresponds to a condition elaborated in [18, Theorem 4.4] combined with the conditions [18, Equation (4.15), (4.21)] stated in the definition of the weighting function φ. We emphasize that a smallness assumption on H has already been necessary for [18, Theorem 4.4]. Thus, condition C can be considered as its counterpart within our analysis. Finally, condition D is due to the variational crimes committed in the advective part B_A. We remark that for $r_{vc,A} = 0$ the conditions used in the above proof are similar to the ones used in [18, Theorem 4.4].

Remark 5.11. Combining estimates (53) and (57), we will show in Section 5.2 that for locally periodic data (cf. Assumption 4.5) satisfying (**H1**), (**H2**) and (**B1**)

$$r_{vc,A} \leq \begin{cases} C\Big(H + \big(\frac{h}{\varepsilon}\big)^{q+1}\Big) & \text{if } W(K_\delta) = W^1_{per}(K_\delta) \text{ and } \frac{\delta}{\varepsilon} \in \mathbb{N} \\ C\Big(H + \big(\frac{\varepsilon}{\delta}\big)^{1/2} + \big(\frac{h}{\varepsilon}\big)^{q+1}\Big) & \text{if } W(K_\delta) = H^1_0(K_\delta) \text{ and } \delta > \varepsilon \end{cases}.$$

Thus, under appropriate coupling conditions and parameters H, h, δ, the term $r_{vc,A}$ can be arbitrarily small and the condition $r_{vc,A} < R_0$ for a $R_0 > 0$ of Theorem 4.3 can be satisfied.

Proof of Theorem 4.4. In order to show the uniform boundedness of B we use the results of Lemma 5.3, Corollary 5.2 and follow the steps of the proof of Lemma 5.9. For $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$, we obtain

$$\big|B_D(v^H, w^H)\big| \leq C |\!|\!| v^H |\!|\!|_D |\!|\!| w^H |\!|\!|_D, \qquad \big|B_A(v^H, w^H)\big| \leq C\,\mathrm{Pe}^{1/2}\, |\!|\!| v^H |\!|\!|_D |\!|\!| w^H |\!|\!|_A,$$

where the constants C are independent of $H, h, \varepsilon, \delta, \alpha$. Combining the uniform boundedness with Theorem 4.3 leads to the existence and uniqueness of the solution of (12) and estimate (22). □

5.2 Proof of the a priori error estimates

In this section, we derive the a priori error estimates for DG-HMM for advection-diffusion problems with locally periodic data (Assumption 4.5). The analysis is performed in two steps: first, in Section 5.2.1, we estimate the macroscopic error of the discontinuous Galerkin method e_{mac} and identify the term e_{HMM} explicitly. Then, in Section 5.2.2, the micro error and the modeling error are estimated. Combining both steps gives Theorem 4.7.

5.2.1 Semi-discrete error

In order to estimate the macro error e_{mac} we introduce a discontinuous Galerkin finite element method for the homogenized problem (4) using numerical integration. This method (never used in practice as the data of (4) are usually not known) is only defined for the convergence analysis.

Single scale method with quadrature. For $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$ we define the bilinear form $\tilde{B}_0$ as the sum $\tilde{B}_0 = \tilde{B}_{D,0} + \tilde{B}_{A,0}$ where $\tilde{B}_{D,0}$ and $\tilde{B}_{A,0}$ are given by

$$\tilde{B}_{D,0}(v^H,w^H) = \sum_{K\in\mathcal{T}_H} |K|\, a^0(x_K)\nabla v^H(x_K)\cdot\nabla w^H(x_K) - \int_\Gamma \{a^0(x_K)\nabla v^H(x_K)\}\cdot[\![w^H]\!]ds + \int_\Gamma \tilde{\mu}_S[\![v^H]\!]\cdot[\![w^H]\!]ds, \tag{48}$$

$$\tilde{B}_{A,0}(v^H,w^H) = \sum_{K\in\mathcal{T}_H} |K|\, b^0(x_K)\cdot\nabla v^H(x_K)\, w^H(x_K) - \int_\Gamma \{b^0(x_K)\}\cdot[\![v^H]\!]w^H_{\tilde{\imath}_0}ds,$$

where for an interior edge $e\subset\overline{K_1}\cap\overline{K_2}$, for some $K_1,K_2\in\mathcal{T}_H$,

$$\{a^0(x_K)\nabla v^H(x_K)\} = \frac{1}{2}\big(a^0(x_{K_1})\nabla v^H(x_{K_1}) + a^0(x_{K_2})\nabla v^H(x_{K_2})\big), \quad \{b^0(x_K)\} = \frac{1}{2}\big(b^0(x_{K_1}) + b^0(x_{K_2})\big).$$

Further, the penalty weighting function $\tilde{\mu}_S$ is given piecewisely by $\tilde{\mu}_S|_e = \big\|\{a^0(x_K)\}\big\|_{\mathcal{F}}\alpha H_e^{-1}$, where the penalization parameter $\alpha>1$ is a positive parameter independent of the local mesh size and the data a^0, and $\tilde{\imath}_0$ denotes the trace taken from the inflow element with respect to $\{b^0(x_K)\}$ - cf. Remark 3.2. Then, we denote by $\tilde{u}^{0,H}$ the solution of the variational problem: find $\tilde{u}^{0,H}\in V^1(\Omega,\mathcal{T}_H)$ such that

$$\tilde{B}_0(\tilde{u}^{0,H},v^H) = \int_\Omega f v^H dx \quad \forall\, v^H\in V^1(\Omega,\mathcal{T}_H). \tag{49}$$

In the proof of Theorem 4.6 we will use the following estimates obtained by combining the interpolation estimates (30) and the scaled trace inequality (32)

$$|\!|\!| u^0 - P_H u^0 |\!|\!|_D \le C a_\infty^{1/2} H \big|u^0\big|_{H^2(\Omega)}, \qquad |\!|\!| u^0 - P_H u^0 |\!|\!|_A \le C b_\infty^{1/2} H^{3/2}\big|u^0\big|_{H^2(\Omega)}, \tag{50}$$

where $P_H u^0\in V^1(\Omega,\mathcal{T}_H)$ denotes the L^2 projection of u^0 onto $V^1(\Omega,\mathcal{T}_H)$ and u^0 is assumed to be in $H^2(\Omega)$. Further, the L^2 projection $P_H u^0$ is bounded in the $|\!|\!|\cdot|\!|\!|$ norm

$$|\!|\!| P_H u^0 |\!|\!|_D \le C a_\infty^{1/2}\big\|u^0\big\|_{H^2(\Omega)}, \qquad |\!|\!| P_H u^0 |\!|\!|_A \le C b_\infty^{1/2}\big\|u^0\big\|_{H^2(\Omega)}, \tag{51}$$

where the bounds on $\big|P_H u^0\big|_{*,\cdot}$ are derived using the fact that $\big|P_H u^0\big|_{*,\cdot} = \big|P_H u^0 - u^0\big|_{*,\cdot}$ as $H^2(\Omega)\hookrightarrow C^0(\Omega)$ for $d\le 3$ and $u^0=0$ on $\partial\Omega$.

The next step is to estimate the difference between B and $\tilde{B}_0$.

Lemma 5.12. *Let $v^H, w^H\in V^1(\Omega,\mathcal{T}_H)$ then*

$$\Big|B_D(v^H,w^H) - \tilde{B}_{D,0}(v^H,w^H)\Big| \le C a_\infty^{1/2} r_{HMM,D}\Big(\big\|\nabla v^H\big\|^2_{L^2(\Omega)} + \big|v^H\big|^2_{*,D}\Big)^{1/2} |\!|\!| w^H |\!|\!|_D,$$

$$\Big|B_A(v^H,w^H) - \tilde{B}_{A,0}(v^H,w^H)\Big| \le C b_\infty^{1/2} r_{HMM,A}\Big(\big\|\nabla v^H\big\|^2_{L^2(\Omega)} + \big|v^H\big|^2_{*,D}\Big)^{1/2} |\!|\!| w^H |\!|\!|_A,$$

where $r_{HMM,D}$ and $r_{HMM,A}$ are defined in (24) *and C is independent of H,h,δ,ε and α.*

Proof. We start by estimating the difference in the diffusive part

$$\begin{aligned}\Big|B_D(v^H,w^H) - \tilde{B}_{D,0}(v^H,w^H)\Big| &\le \left|\sum_{K\in\mathcal{T}_H}|K|\,(a^0_K - a^0(x_K))\nabla v^H(x_K)\cdot\nabla w^H(x_K)\right| + \left|\int_\Gamma\{\big(a^0_K - a^0(x_K)\big)\nabla v^H(x_K)\}\cdot[\![w^H]\!]ds\right| \\ &\quad + \left|\sum_{e\in\mathcal{E}}\big(\big\|\{a^0_K\}\big\|_{\mathcal{F}} - \big\|\{a^0(x_K)\}\big\|_{\mathcal{F}}\big)\int_e \alpha H_e^{-1}[\![v^H]\!]\cdot[\![w^H]\!]ds\right| = |I_1| + |I_2| + |I_3|.\end{aligned}$$

The terms I_1 and I_2 are estimated by following the proof of [7, Lemma 5.10]

$$|I_1| \le a_\infty r_{HMM,D}\big\|\nabla v^H\big\|_{L^2(\Omega)}\big\|\nabla w^H\big\|_{L^2(\Omega)}, \qquad |I_2| \le C_g\alpha^{-1/2}a_\infty r_{HMM,D}\big\|\nabla v^H\big\|_{L^2(\Omega)}\big|w^H\big|_{*,D},$$

and I_3 can be bounded by using the reverse triangle inequality

$$|I_3| \le \int_\Gamma \big\|\{a^0_K\} - \{a^0(x_K)\}\big\|_{\mathcal{F}}\alpha H_e^{-1}[\![v^H]\!]\cdot[\![w^H]\!]ds \le a_\infty r_{HMM,D}\big|v^H\big|_{*,D}\big|w^H\big|_{*,D}.$$

Further, the estimate for the advective part can be derived straightforwardly from Lemma 5.7. □

Proof of Theorem 4.6. We first split the error into $|||u^0 - u^H||| \leq |||u^0 - P_H u^0||| + |||P_H u^0 - u^H|||$ using the L^2 projection $P_H u^0$. Then, the stability of the multiscale method allows us to estimate the second term

$$\begin{aligned}
\alpha_S |||P_H u^0 - u^H||| &\leq \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{B(P_H u^0 - u^H, w^H)}{|||w^H|||} = \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{B(P_H u^0, w^H) - \int_\Omega f w^H dx}{|||w^H|||} \\
&= \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{B(P_H u^0, w^H) - \tilde{B}_0(\tilde{u}^{0,H}, w^H)}{|||w^H|||} \\
&= \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{B(P_H u^0, w^H) - \tilde{B}_0(P_H u^0, w^H) + \tilde{B}_0(P_H u^0 - \tilde{u}^{0,H}, w^H)}{|||w^H|||}.
\end{aligned}$$

Hence, we get for the error

$$\begin{aligned}
|||u^0 - u^H||| \leq |||u^0 - P_H u^0||| &+ \frac{1}{\alpha_S} \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{\tilde{B}_0(P_H u^0 - \tilde{u}^{0,H}, w^H)}{|||w^H|||} \\
&+ \frac{1}{\alpha_S} \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{B(P_H u^0, w^H) - \tilde{B}_0(P_H u^0, w^H)}{|||w^H|||}.
\end{aligned}$$

In view of (23) the first two terms quantify the macroscopic error e_{mac} and the third term representing e_{HMM}, quantifies the error due to our multiscale strategy. The first term can be estimated using (50), while the second term (including the contribution to the macroscopic error arising from the use of numerical quadrature) is estimated in Theorem A.1. Finally, Lemma 5.12 allows to the estimate the third term

$$e_{HMM} \leq C\Big(a_\infty^{1/2} r_{HMM,D} + b_\infty^{1/2} r_{HMM,A}\Big)\Big(\big\|\nabla P_H u^0\big\|^2_{L^2(\Omega)} + \big|P_H u^0\big|^2_{*,D}\Big)^{1/2} |||w^H|||.$$

The approximation property (50) and the boundedness (51) of the L^2 projection conclude the proof. □

5.2.2 Fully discrete error

In this section, we first estimate the contribution of the micro error $r_{mic,D}$ and $r_{mic,A}$ defined in (25) and (26), respectively (due to the standard FEM on the micro mesh) for general oscillating data a^ε, b^ε. In a second step, we derive bounds for the modeling error $r_{mod,D}$, $r_{mod,A}$ (see again (25) and (26)) for locally periodic data (see Assumption 4.5). Finally, we combine the estimates for macro, micro and modeling error to prove Theorem 4.7.

Micro error. Important ingredients for the a priori estimates for the micro errors $r_{mic,D}$ and $r_{mic,A}$ are Assumption (**H1**) and (**B1**) providing the necessary regularity of ψ_K^i as well as the proper scaling with respect to ε of the derivatives of ψ_K^i and b^ε. We emphasize that Lemma 5.13 is valid for data a^ε, b^ε without any assumption on their spatial structure.

Lemma 5.13. *Consider the micro finite element space defined in* (7) *with* $q \in \mathbb{N}_{>0}$ *and assume that* (2), ***(H1)*** *and* ***(B1)***. *Furthermore, we assume that Dirichlet boundary conditions are used in* (13) *and* (14) *for general data* $a^\varepsilon(x)$, $b^\varepsilon(x)$. *For the special case when* $a^\varepsilon(x) = a(x_K, x/\varepsilon) = a(x_K, y)$ *and* $b^\varepsilon(x) = b(x_K, x/\varepsilon) = b(x_K, y)$ *are* Y*-periodic in* y, *collocated in the slow variables at the quadrature points* x_K *of the sampling domain* K_δ *and* $\delta/\varepsilon \in \mathbb{N}$, *we assume that periodic boundary conditions are used in* (13) *and* (14). *Then, for any* $K \in \mathcal{T}_H$,

$$\big\|\overline{a}_K^0 - a_K^0\big\|_{\mathcal{F}} \leq C a_\infty \left(\frac{h}{\varepsilon}\right)^{2q}, \tag{52}$$

$$\Big|\overline{b}_K^0 - b_K^0\Big| \leq C b_\infty \left(\frac{h}{\varepsilon}\right)^{q+1}, \tag{53}$$

where C *is independent of* H, h, ε *and* δ.

Proof. The error estimate (52) for non-symmetric tensors a^ε has been derived in [27] and [12, Lemma 4.6] (for symmetric tensors the error bound has first been published in [3, 4] for piecewise linear micro functions and stated for higher order micro functions in [7, Lemma 5.2]). We thus proceed with proving (53). Let $1 \leq j \leq d$ and $K \in \mathcal{T}_H$. Integrating by parts leads to

$$\int_{K_\delta} b^\varepsilon \cdot \nabla \psi_K^{j,h} dx = \int_{\partial K_\delta} b^\varepsilon \psi_K^{j,h} \cdot n \, ds - \int_{K_\delta} \operatorname{div}(b^\varepsilon) \psi_K^{j,h} dx = - \int_{K_\delta} \operatorname{div}(b^\varepsilon) \psi_K^{j,h} dx,$$

where the boundary integral vanishes. Indeed, for periodic data and $\delta/\varepsilon \in \mathbb{N}$ we choose periodic coupling in (8) and we observe that $b^\varepsilon(x)\psi_K^{j,h}(x)$ is K_δ-periodic. Therefore, the values on opposing faces cancel. In all other cases, we choose Dirichlet coupling in (8) and thus $\psi_K^{j,h}(x)$ is equal to zero on the boundary. Analogously, it holds $\int_{K_\delta} b^\varepsilon \cdot \nabla \psi_K^j dx = -\int_{K_\delta} \operatorname{div}(b^\varepsilon)\psi_K^j dx$. Hence, we examine the j-th entry of the difference of the vectors b_K^0 and $\overline{b}_K^0$

$$(b_K^0 - \overline{b}_K^0)_j = \frac{1}{|K_\delta|}\int_{K_\delta} b^\varepsilon(x)\cdot(\nabla\psi_K^{j,h} - \nabla\psi_K^j)dx = \frac{1}{|K_\delta|}\int_{K_\delta} \operatorname{div}(b^\varepsilon(x))\Big(\psi_K^j - \psi_K^{j,h}\Big)dx.$$

Using standard finite element result for the L^2 error, Assumptions (**H1**) and (**B1**) leads to

$$\begin{aligned}\Big|(b_K^0 - \overline{b}_K^0)_j\Big| &\le \frac{1}{|K_\delta|}\left(\int_{K_\delta} |\operatorname{div} b^\varepsilon|^2 dx\right)^{1/2}\Big\|\psi_K^j - \psi_K^{j,h}\Big\|_{L^2(K_\delta)} \\ &\le C\frac{1}{|K_\delta|}|b^\varepsilon|_{W^{1,\infty}(K_\delta)}\sqrt{|K_\delta|}h^{q+1}\Big|\psi_K^j\Big|_{H^{q+1}(K_\delta)} \le Cb_\infty\left(\frac{h}{\varepsilon}\right)^{q+1}. \qquad \square\end{aligned}$$

Remark 5.14. The optimality of the estimate of the micro error due to advection (53) is shown for $q = 1$ in Section 6. In contrast, this is an open issue for $q > 1$. Further, if b^ε is a linear combination of the rows of a^ε, i.e., there exist $\beta_1, \dots, \beta_d \in \mathbb{R}$ independent of x such that $(b^\varepsilon(x))^T = \sum_{i=1}^d \beta_i a_i^\varepsilon(x)$, following [7, Lemma 5.2], one can show that the micro error due to advection can again be estimated by $Cb_\infty(\frac{h}{\varepsilon})^{2q}$.

Modeling error. Next, we provide estimates for the modeling error $r_{mod,D}$ and $r_{mod,A}$. We remark, that for data a^ε, b^ε without any assumptions about their spatial structure, the modeling error cannot be estimated in general. Thus, the estimates of Lemma 5.15 are derived for locally periodic data.

For the class of locally periodic data results in periodic homogenization theory show, that the whole sequence $\{u^\varepsilon\}$ weakly converges to u^0, the solution of the homogenized problem (4) (see [34, p. 31], [19, Sect. 13]). Furthermore, the homogenized quantities $a^0(x)$ and $b^0(x)$ can be calculated as special averages involving the solutions of (infinitely many) cell problems. For an arbitrary $x \in \Omega$, let us introduce the first order correctors $\chi^k(x,y) : \Omega \times Y \to \mathbb{R}$ with $\chi^k(x,\cdot) \in W^1_{per}(Y)$, for $k = 1, \dots, d$, the unique solution of the cell problem

$$\int_Y a(x,y)\nabla\chi^k(x,y)\cdot\nabla v(y)dy = -\int_Y a(x,y)e_k\cdot\nabla v(y)dy \qquad \forall\, v \in W^1_{per}(Y).$$

Then, the homogenized tensor $a^0(x)$ at $x \in \Omega$ is given by

$$a^0_{ij}(x) = \frac{1}{|Y|}\int_Y a_{ij}(x,y) + \sum_{k=1}^d a_{ik}(x,y)\frac{\partial\chi^j}{\partial y_k}(x,y)dy, \quad 1 \le i,j \le d, \tag{54}$$

and the homogenized velocity field b^0 has the explicit representation

$$b^0_j(x) = \frac{1}{|Y|}\int_Y b_j(x,y) + \sum_{k=1}^d b_k(x,y)\frac{\partial\chi^j}{\partial y_k}(x,y)dy, \quad 1 \le j \le d. \tag{55}$$

Based on the representations (54) and (55), we can estimate the modeling error.

Lemma 5.15. *Let $K \in \mathcal{T}_H$. We assume that the data $a^\varepsilon(x)$, $b^\varepsilon(x)$ satisfy (2) and Assumption 4.5, and are collocated in the slow variable, i.e., $a^\varepsilon(x) = a(x_K, x/\varepsilon)$, $b^\varepsilon(x) = b(x_K, x/\varepsilon)$, where x_K is the quadrature node of the sampling domain K_δ. Then*

$$\big\|a^0(x_K) - \overline{a}^0_K\big\|_{\mathcal{F}} \le \begin{cases} 0 & \text{if } W(K_\delta) = W^1_{per}(K_\delta) \text{ and } \frac{\delta}{\varepsilon} \in \mathbb{N} \\ Ca_\infty\frac{\varepsilon}{\delta} & \text{if } W(K_\delta) = H^1_0(K_\delta) \text{ and } \delta > \varepsilon \end{cases}, \tag{56}$$

$$\Big|b^0(x_K) - \overline{b}^0_K\Big| \le \begin{cases} 0 & \text{if } W(K_\delta) = W^1_{per}(K_\delta) \text{ and } \frac{\delta}{\varepsilon} \in \mathbb{N} \\ Cb_\infty\big(\frac{\varepsilon}{\delta}\big)^{1/2} & \text{if } W(K_\delta) = H^1_0(K_\delta) \text{ and } \delta > \varepsilon \end{cases}. \tag{57}$$

where C is independent of ε, δ.

Proof. The estimates (56) for the modeling error due to diffusion have been derived in [29, Theorem 3.2] and in [6, Proposition 14] (for the case $\delta/\varepsilon \in \mathbb{N}$). Therefore, we discuss the proof of estimates (57). The first estimate is derived by following the proof of [6, Proposition 14]. The second estimate follows [29, Theorem 3.2] and [34, Equation (1.51)]. □

Proof of Theorem 4.7. Follows from the combination of Theorem 4.6 and Lemma 5.13. □

Remark 5.16. We notice that without Assumption (**H2**), i.e., collocating $a^\varepsilon(x)$ and $b^\varepsilon(x)$ in the slow variable x, additional terms of order δ arise in estimates (56) and (57). Further, in view of Lemma 5.15, the coupling $\delta/\varepsilon \in \mathbb{N}$ and periodic boundary conditions is optimal for locally periodic data. Finally, we remark, that for a one-dimensional homogenization problem with periodic data a resonance error due to advection of order ε/δ can be shown. Thus, the optimality of estimate (57) is an open question.

6 Numerical results

In this section, we present numerical experiments confirming the sharpness of the convergence rates derived in Theorem 4.7. Furthermore, we apply the proposed method on an advection dominated problem where the solution exhibits a boundary layer and we illustrate the applicability of our multiscale strategy for nonperiodic (random) tensors with variable cell size.

6.1 Convergence rates

In order to corroborate our theoretical convergence rates, we first choose a simple periodic problem with known homogenized data a^0 and b^0 and analytically known homogenized solution u^0. If we choose periodic boundary conditions for the micro problems (8) and the size δ of the sampling domains K_δ such that $\delta/\varepsilon \in \mathbb{N}$, the modeling error vanishes and we can verify numerically the macro and micro convergence rates. Then, for piecewise linear micro elements, i.e., $q = 1$, the error in the $|||\cdot|||$ norm satisfies

$$|||u^0 - u^H||| \leq C\left(a_\infty^{1/2} H + b_\infty^{1/2} H^{3/2} + a_\infty^{1/2}\left(\frac{h}{\varepsilon}\right)^2 + b_\infty^{1/2}\left(\frac{h}{\varepsilon}\right)^2\right),$$

which is a robust convergence rate (i.e., independent of ε). We emphasize that for periodic data the use of numerical integration on the macro scale does not have any influence as in this case the homogenized data are constant. Thus we can omit the term $b_\infty^{1/2} \min\{\mathrm{Pe}^{1/2} H^2, H\}$ in estimate (27).

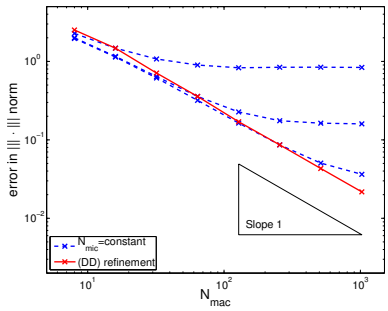
We consider problem (1) on the domain $\Omega = (0,1)^2$ with a tensor $a^\varepsilon(x) = \nu\, \tilde{a}^\varepsilon(x)$ and a velocity field $b^\varepsilon(x) = \Phi(x/\varepsilon)(1,1)^T$, where $\tilde{a}^\varepsilon(x) = \Phi(x/\varepsilon) Id$, $\nu > 0$, and

$$\Phi(y) = \frac{64}{9\sqrt{17}}\left(\sin(2\pi y_1) + \frac{9}{8}\right)\left(\cos(2\pi y_2) + \frac{9}{8}\right).$$

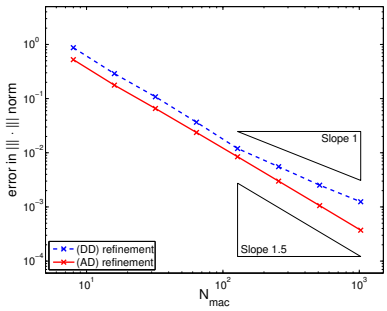
The homogenized quantities a^0 and b^0 are given by $a^0 = \nu\, Id$ and $b^0 = (1,1)^T$ (cf. [34, Chapter 1.2] and Remark 5.14) leading to $\mathrm{Pe} = \operatorname{diam}\Omega/\nu$. The source f is adjusted such that the homogenized solution u^0 is given by $u^0(x) = \sin(2\pi x_1)\sin(2\pi x_2)$. Further, we choose the size of the sampling domains $\delta = \varepsilon$ and we investigate a diffusion and an advection dominated problem with $\mathrm{Pe} = 1$ and $\mathrm{Pe} = 10^6$, respectively.

We consider a family of macro partitions $\mathcal{T}_H$ with $2 \cdot N_{mac}^2$ triangles, which are generated by uniformly refining a coarse mesh $\widetilde{\mathcal{T}}$ (N_{mac} denotes the number of macro elements in each spatial dimension). The initial macro partition $\widetilde{\mathcal{T}}$ in turn is constructed by randomly perturbing a uniform mesh with $N_{mac} = 8$. Therefore, the ratio H/N_{mac} is constant for different $\mathcal{T}_H$. Such non-uniform meshes are chosen in order to prevent possible super-convergence for meshes suitably aligned with the velocity field (cf. [24]). Further, we choose $\alpha = 10$ as penalization parameter for the diffusive part B_D.

In Figure 1 the error between u^0 and the DG-HMM numerical solution in the $|||\cdot|||$ norm is plotted under different refinement schemes for N_{mac} and N_{mic} (described in Section 4.2). In Figure 1.(a) we observe that without simultaneous refinement of H and h/ε the micro error becomes dominant for large N_{mac} leading to an overall error independent of N_{mac}. The optimality of the simultaneous refinement strategy (DD) for purley diffusive problems has been discussed in [6]. In Figure 1.(b) we employ the simultaneous refinement strategies (DD) and (AD) for the advection dominated problem. We emphasize that the refinement strategy (DD) is not sufficient to obtain the rate 1.5 for the advection dominated problem as the micro error converging with a linear rate with respect to H becomes dominant for $N_{mac} > 100$. In summary, we



(a) Diffusion dominated problem with $\mathrm{Pe} = 1$. Optimal simultaneous refinement (DD) (solid line). Constant micro mesh $N_{mic} = 4, 8, 16$ (dashed lines).



(b) Advection dominated problem with $\mathrm{Pe} = 10^6$. Optimal simultaneous refinement (AD) (solid line). Simultaneous refinement (DD) (dashed line).

Figure 1: Homogenization test problem of Section 6.1 with periodic data. Error in $|||\cdot|||$ norm as a function of N_{mac}. Macro meshes with $N_{mac} = 8, 16, 32, 64, 128, 256, 512, 1024$. Refinement strategy (DD) with $N_{mic} \sim \sqrt{N_{mac}}$ where $N_{mic} = 3, 4, 6, 8, 12, 16, 23, 32$. Refinement strategy (AD) with $N_{mic} \sim (N_{mac})^{3/4}$ where $N_{mic} = 5, 8, 14, 23, 39, 64, 108, 182$.

observe that the convergence rates predicted by Theorem 4.7 are verified numerically for problems with constant homogenized data.

As next step, we modify the precedent test problem by taking locally periodic data. We replace $\tilde{a}^\varepsilon(x)$ and $b^\varepsilon(x)$ by (cf. Examples of [8, Section 5])

$$\tilde{a}^\varepsilon(x) = \begin{pmatrix} \Psi_1\big(x, \frac{x}{\varepsilon}\big) & 0 \\ 0 & \Psi_2\big(x, \frac{x}{\varepsilon}\big) \end{pmatrix}, \qquad b^\varepsilon(x) = \begin{pmatrix} \Psi_1\big(x, \frac{x}{\varepsilon}\big) \\ \Psi_2\big(x, \frac{x}{\varepsilon}\big) \end{pmatrix},$$

$$\Psi_1(x,y) = \left(x_1^3 + 3 + \frac{2\sqrt{17}}{8\sin(2\pi y_1) + 9}\right)^{-1}, \qquad \Psi_2(x,y) = \left(x_2^2 + 0.05 + (x_1x_2 + 1)\frac{2\sqrt{17}}{8\cos(2\pi y_2) + 9}\right)^{-1}.$$

Hence, $a^\varepsilon(x)$ is an anisotropic, locally periodic tensor. The data is chosen such that the homogenized quantities are given by the elementwise harmonic mean

$$a^0(x) = \nu \begin{pmatrix} \Psi_1^0(x) & 0 \\ 0 & \Psi_2^0(x) \end{pmatrix}, \qquad b^0(x) = \begin{pmatrix} \Psi_1^0(x) \\ \Psi_2^0(x) \end{pmatrix},$$

$$\Psi_1^0(x) = \big(x_1^3 + 5\big)^{-1}, \qquad \Psi_2^0(x,y) = \big(x_2^2 + 0.05 + 2(x_1x_2 + 1)\big)^{-1},$$

where we remark that $-\operatorname{div} b^0(x) \geq 0$ holds on Ω. Beside the different data a^ε, b^ε we exactly take the same setting as in the precedent test for periodic data (here we also adjust f such that $u^0(x) = \sin(2\pi x_1)\sin(2\pi x_2)$). Additionally, for solving the micro problem (8) on an element $K \in \mathcal{T}_H$, we collocate Ψ_1 and Ψ_2 in the slow variable x at the barycenter x_K.

By comparing the results of Figure 2 for locally periodic data to the results of Figure 1 we identify the same behavior for both advection and diffusion dominated problems under different refinement strategies. We observe that the effect of the numerical integration in the advective part does not alter the convergence rate for this test case (see Remark 4.8).

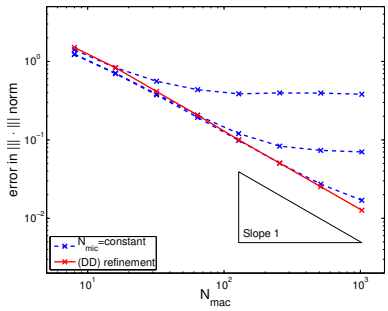
We emphasize that for the advection dominated problem the evaluation of the right-hand side $\int_\Omega f v^H dx$ of the variational problem (12) is done using a high order quadrature formula in order to overkill the effect of numerical integration in the evaluation of the right-hand side $\int_\Omega f\, v^H dx$.

6.2 Advection dominated multiscale problem with a boundary layer

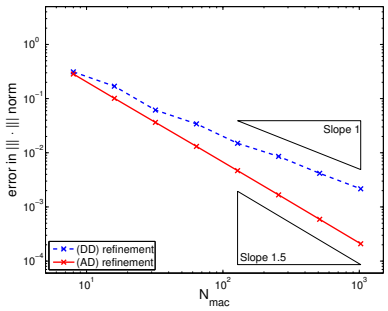
We consider the periodic data $a^\varepsilon, b^\varepsilon$ used in the first part of Section 6.1 and adjust f such that the homogenized solution u^0 is given by

$$u^0(x) = x_1x_2\left(1 + \frac{e^{-\mathrm{Pe}} - e^{-\mathrm{Pe}(1-x_1)(1-x_2)}}{1 - e^{-\mathrm{Pe}}}\right),$$

which exhibits a boundary layer of width $\mathcal{O}(\mathrm{Pe}^{-1})$ near to $\{1\} \times [0,1] \cup [0,1] \times \{1\}$. Such problems have been used as model problems for single scale methods (see [18, Example 4] and the references therein). We



(a) Diffusion dominated problem with Pe = 1. Optimal simultaneous refinement (DD) (solid line). Constant micro mesh $N_{mic} = 4, 8, 16$ (dashed lines).



(b) Advection dominated problem with Pe = 10^6. Optimal simultaneous refinement (AD) (solid line). Simultaneous refinement (DD) (dashed line).

Figure 2: Homogenization test problem of Section 6.1 with locally periodic data. Error in $|||\cdot|||$ norm as a function of N_{mac}. Macro meshes with $N_{mac} = 8, 16, 32, 64, 128, 256, 512, 1024$. Refinement strategy (DD) with $N_{mic} \sim \sqrt{N_{mac}}$ where $N_{mic} = 3, 4, 6, 8, 12, 16, 23, 32$. Refinement strategy (AD) with $N_{mic} \sim (N_{mac})^{3/4}$ where $N_{mic} = 5, 8, 14, 23, 39, 64, 108, 182$.

compare the behavior of the two multiscale methods DG-HMM, described in this article, and FE-HMM with a macro solver based on standard FEM. The FE-HMM is built on the method described in [11] by adding the advective part $\sum_{K\in\mathcal{T}_H} \frac{|K|}{|K_\delta|} \int_{K_\delta} b^\varepsilon \cdot \nabla v_K^h \, w^H(x_K) dx$, the first term of B_A defined in (6).

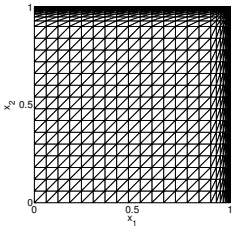
We choose the size of the sampling domains $\delta = \varepsilon$, periodic coupling in the micro problems (8), the penalization parameter $\alpha = 10$ and $N_{mic} = 16$, the number of micro elements in each spatial dimension in the sampling domains. Further, we construct a highly anisotropic mesh consisting of 1800 triangles (see Figure 3.(a)) such that the boundary layer for Pe = 10^3 is properly resolved. We can observe in Figure 3.(b) and Figure 3.(e) that on this macro triangulation both numerical methods, DG-HMM and FE-HMM, are able to capture the boundary layer for Pe = 10^3 correctly and produce similar results. Next, we investigate the robustness of the two methods with respect to the Péclet number Pe. We take the macro triangulation given in Figure 3.(a) and compute the numerical solution using DG-HMM and FE-HMM for Pe $\in \{10^3, \ldots, 10^8\}$. We see in Figure 3.(f) that for Pe = 10^5 the result obtained by the FE-HMM exhibits unphysical oscillations polluting the numerical solution on the entire domain Ω due to the standard FEM macro solver. In contrast, as DG-HMM is based on a discontinuous Galerkin macro solver, it produces qualitatively good results apart from oscillations limited to the boundary layer (see Figure 3.(c)). This nice robustness of the DG-HMM with respect to the Péclet number Pe can as well be observed in Figure 3.(d) where we compare the error in the L^2 norm for DG-HMM and FE-HMM for different Péclet numbers Pe on the macro mesh given in Figure 3.(a). While the error for the FE-HMM depends linearly on Pe, the error for the DG-HMM only slightly increases due to the peaks in the boundary layer. We emphasize that such results for DG-HMM can only be obtained for a penalization term μ which scales with the magnitude of the diffusion tensor a^0 which is achieved by including the factor $\left\|\{a_K^0\}\right\|_{\mathcal{F}}$ in (6).

6.3 Example with non-periodic, random data

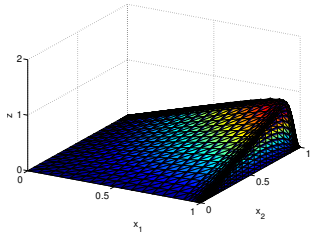
In practice, the diffusion tensor a^ε and the velocity field b^ε may neither be known analytically nor satisfy a periodicity assumption. In porous media, log-normal fields are often used to model the heterogeneities of the media (e.g., see [39]). We consider problem (1) with a random tensor a^ε and a random velocity field b^ε based on log-normal stochastic fields $W_a(x)$ and $W_b(x)$ on Ω with an underlying normal distribution of mean zero and variance $\sigma^2 = 0.5$ (cf. [11, Section 4.2]). The random fields are generated by a moving ellipse average method with correlation lengths $\varepsilon_x = 0.01, \varepsilon_y = 0.02$ at 5000^2 discrete points. For an arbitrary $x \in \Omega$ we use bilinear interpolation. We set $a^\varepsilon(x) = W_a(x)\, Id$, $b^\varepsilon(x) = W_b(x)\,(1,0)^T$ and $f(x) = 1$ on Ω.

First, we compute a reference solution using a standard FEM on a fine mesh with 10^6 degrees of freedom (Figure 4.(b)). Then, we apply DG-HMM for a uniform macro mesh of 2048 triangles and we take again the penalization parameter $\alpha = 10$. The numerical solution is computed for sampling domains of different sizes δ_i, $i = 1, \ldots, 5$. Additionally, $N_{mic,i}$ is chosen such that the micro mesh size $h = \delta_i / N_{micro_i}$ remains constant. Finally, we employ Dirichlet coupling for the micro problems (8).

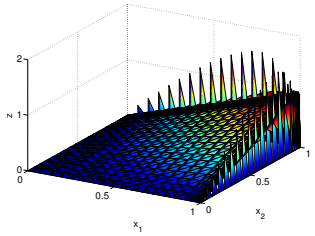
We notice that the fine scale solution approximates u^ε and some care is needed in order to compare



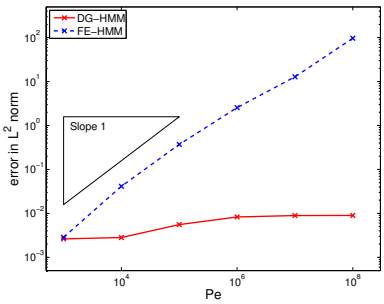
(a) Anisotropic mesh with 1800 triangles resolving boundary layer for $\mathrm{Pe} = 10^3$.



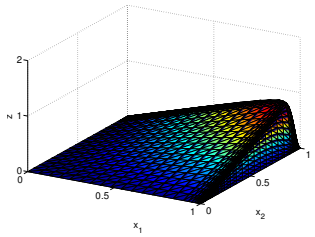
(b) DG-HMM. $\mathrm{Pe} = 10^3$.



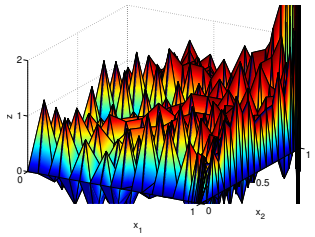
(c) DG-HMM. $\mathrm{Pe} = 10^5$.



(d) Error in L^2 norm as a function of Pe.

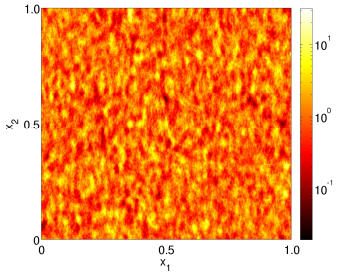


(e) FE-HMM. $\mathrm{Pe} = 10^3$.

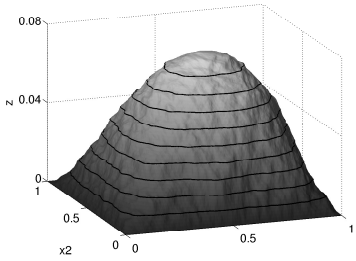


(f) FE-HMM. $\mathrm{Pe} = 10^5$.

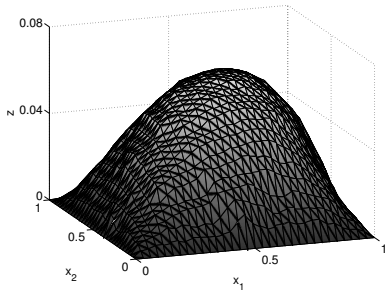
Figure 3: Advection dominated test problem with boundary layer of Section 6.2 with $\mathrm{Pe} \in \{10^3, 10^5\}$. DG-HMM based on DG-FEM macro solver. FE-HMM based on standard FEM macro solver. Error in L^2 norm as a function of Pe for DG-HMM and FE-HMM on an anisotropic mesh.



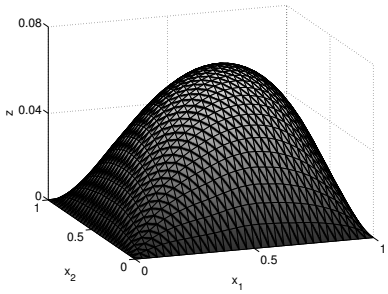
(a) $W_a(x_1, x_2)$.



(b) Standard FEM. 10^6 degrees of freedom.



(c) DG-HMM. $N_{mac} = 32, \delta_1 = 0.015,\ N_{mic,1} = 8$.



(d) DG-HMM. $N_{mac} = 32, \delta_5 = 0.24,\ N_{mic,5} = 128$.

Figure 4: Test problem with random data of Section 6.3. Reference fine scale solution based on standard FEM. DG-HMM numerical solutions for different sampling domains.

	$\delta_1 = 0.015$ $N_{mic,1} = 8$	$\delta_2 = 0.03$ $N_{mic,2} = 16$	$\delta_3 = 0.06$ $N_{mic,3} = 32$	$\delta_4 = 0.12$ $N_{mic,4} = 64$	$\delta_5 = 0.24$ $N_{mic,5} = 128$	fine scale
energy norm $\|\cdot\|_E$	0.1768	0.1779	0.1793	0.1811	0.1823	0.1859
error in L^2 norm	0.0032	0.0032	0.0029	0.0023	0.0022	-

Table 1: Energy norm $\|\cdot\|_E$ of the DG-HMM and fine scale solutions (standard FEM), error in L^2 norm between the DG-HMM solutions and the fine scale solution (standard FEM) for sampling domains of different sizes δ_i for the test problem of Section 6.3.

the fine scale and DG-HMM solutions. It is already known for pure diffusion problems that the error in the L^2 norm between the DG-HMM solutions and the fine scale solution can be $\mathcal{O}(\varepsilon)$ close, while the error measured in the H^1 norm is $\mathcal{O}(1)$. Therefore, we choose to compare the energy norm $\|\cdot\|_E$ rather than to compute the error in the H^1 norm.

In Table 1 we compare the (broken) energy norm $\|u\|_E = \sqrt{\sum_{K\in\mathcal{T}_H} \int_K a_K^0 \nabla u \cdot \nabla u \, dx}$ of the DG-HMM solutions for different sampling domain sizes δ_i to the $\|\cdot\|_E$ norm of the fine scale solution. Further, the error in the L^2 norm between the DG-HMM solutions for different δ_i and the fine scale solution are computed. We observe in Table 1 (see also Figure 4) that for both, the energy norm $\|\cdot\|_E$ and the error in the L^2 norm, improved results are obtained for larger sampling domains.

7 Conclusion

In this paper we have constructed and analyzed a discontinuous Galerkin FE-HMM method for advection-diffusion problems. The method is constructed to allow for general microstructures (not necessarily periodic) and stability results have been established for a general class of advection-diffusion problems for which we relax the usual coercivity condition relating advection and reaction. We also allow for variable diffusion tensors and velocity fields and our analysis includes the cases of advection or diffusion dominated problems. The complexity of our method has been shown to be independent of the smallest scale in the medium and the numerical work scales with the macroscopic degrees of freedom. A priori error estimates in the H^1 norm and convergence to the homogenized solution are proved under the assumption of locally periodic data. To this end, we derived new results about the effect of numerical integration for single scale DG-FEM for advection-diffusion problems. Numerical experiments given for both periodic and non-periodic data, show the capabilities of the proposed method. Finally, an extension of the presented results to meshes with hanging nodes is straightforward, whereas the generalization to higher order macro finite elements in the spirit of [7, Section 5.4] has not been adressed yet at all.

Acknowledgements. This work was supported in part by the Swiss National Science Foundation under Grant 200021 134716/1.

References

[1] A. Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput.*, 23(6):2041–2054, 2002.

[2] A. Abdulle. Multiscale methods for advection-diffusion problems. *Discrete Contin. Dyn. Syst.*, 2005:11–21, 2005.

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

[4] A. Abdulle. Analysis of a heterogeneous multiscale FEM for problems in elasticity. *Math. Models Methods Appl. Sci.*, 16(4):615–635, 2006.

[5] A. Abdulle. Multiscale method based on discontinuous Galerkin methods for homogenization problems. *C. R. Acad. Sci. Paris*, Ser. I 346(1-2):97–102, 2008.

[6] A. Abdulle. The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs. *GAKUTO Internat. Ser. Math. Sci. Appl.*, 31:135–184, 2009.

[7] A. Abdulle. Discontinuous Galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales. *Math. Comp.*, 81(278):687–713, 2012.

[8] A. Abdulle and Y. Bai. Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems. *J. Comput. Phys.*, 231(21):7014–7036, 2012.

[9] A. Abdulle, W. E, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012.

[10] A. Abdulle and A.A. Medovikov. Second order Chebyshev methods based on orthogonal polynomials. *Numer. Math.*, 90(1):1–18, 2001.

[11] A. Abdulle and A. Nonnenmacher. A short and versatile finite element multiscale code for homogenization problems. *Comput. Methods Appl. Mech. Engrg.*, 198(37-40):2839–2859, 2009.

[12] A. Abdulle and G. Vilmart. Analysis of the finite element heterogeneous multiscale method for quasilinear elliptic homogenization problems. *Accepted in Math. Comp.*, 2012.

[13] E. Acerbi. Convergence of second order elliptic operators in complete form. *Boll. Unione Mat. Ital. (B)*, 18(5):539–555, 1981.

[14] S. Agmon. *Lectures on elliptic boundary value problems.* D. Van Nostrand Co., Princeton, NJ - Toronto - London, 1965.

[15] G. Allaire and A. Raphael. Homogenization of a convection-diffusion model with reaction in a porous medium. *C. R. Math. Acad. Sci. Paris*, 344(8):523–528, 2007.

[16] D.N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.

[17] 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(5):1749–1779, 2002.

[18] B. Ayuso and L.D. Marini. Discontinuous Galerkin methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 47(2):1391–1420, 2009.

[19] A. Bensoussan, J.L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*, volume 5. North Holland, 1978.

[20] E. Burman and P. Zunino. A domain decomposition method based on weighted interior penalties for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 44(4):1612–1638, 2006.

[21] S. Chen, W. E, and C.-W. Shu. The heterogeneous multiscale method based on the discontinuous Galerkin method for hyperbolic and parabolic problems. *Multiscale Model. Simul.*, 3(4):871–894, 2005.

[22] P.G. Ciarlet. *The finite element method for elliptic problems*, volume 4. North-Holland, 1978.

[23] P.G. Ciarlet and P.A. Raviart. The combined effect of curved boundaries and numerical integration in isoparametric finite element methods. *In A.K. Aziz (Ed), Math. Foundation of the FEM with Applications to PDE, Academic Press, New York, NY*, pages 409–474, 1972.

[24] B. Cockburn, B. Dong, J. Guzmán, and J. Qian. Optimal convergence of the original DG method on special meshes for variable transport velocity. *SIAM J. Numer. Anal.*, 48(1):133–146, 2010.

[25] C. Dawson, S. Sun, and M.F. Wheeler. Compatible algorithms for coupled flow and transport. *Comput. Methods Appl. Mech. Engrg.*, 193(23-26):2565–2580, 2004.

[26] E. De Giorgi and S. Spagnolo. Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine. *Boll. Unione Mat. Ital.*, 4(8):391–411, 1973.

[27] R. Du and P. Ming. Heterogeneous multiscale finite element method with novel numerical integration schemes. *Commun. Math. Sci.*, 8(4):863–885, 2010.

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

[29] W. E, P. Ming, and P. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.*, 18(1):121–156, 2004.

[30] P. Henning and M. Ohlberger. The heterogeneous multiscale finite element method for advection-diffusion problems with rapidly oscillating coefficients and large expected drift. *Netw. Heterog. Media*, 5(4):711–744, 2010.

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

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

[33] P. Houston, C. Schwab, and E. Süli. Discontinuous hp-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, 39(6):2133–2163, 2002.

[34] V.V. Jikov, S.M. Kozlov, and O.A. Oleinik. *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, 1994.

[35] C. Johnson and J. Pitkäranta. An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation. *Math. Comput.*, 46(173):1–26, 1986.

[36] P. Marcellini. Convergence of second order linear elliptic operators. *Boll. Unione Mat. Ital. (B)*, 16(5):278–290, 1979.

[37] F. Murat and L. Tartar. H-convergence, topics in the mathematical modeling of composite materials. *Progr. Nonlinear Differential Equations Appl.*, (31):21–43, 1997.

[38] M. Stynes. Steady-state convection-diffusion problems. *Acta Numer.*, 14(82):445–508, 2005.

[39] T.C. Wallstrom, S. Hou, M.A. Christie, L.J. Durlofsky, and D.H. Sharp. Accurate scale up of two phase flow using renormalization and nonuniform coarsening. *Comput. Geosci.*, 3(1):69–87, 1999.

[40] T. Yu and X. Yue. Heterogeneous multiscale discontinuous Galerkin method for convection-diffusion problems. *Preprint*, 2011.

A The effect of numerical integration for single scale DG-FEM

In this section, we study the influence of numerical integration for a single scale discontinuous Galerkin method. Without loss of generality we take the homogenized problem (4) as model problem for a single scale advection-diffusion problem. The single scale analysis presented in this section consists of two parts. First, we briefly comment on the analysis of the single scale DG-FEM without numerical quadrature used here for advection-diffusion problems, as it slightly differs from the method analyzed in [18] due to the choice of a different model problem. Second, we derive the stability and a priori results for the single scale DG-FEM with numerical quadrature defined in (49).

A.1 DG-FEM without numerical quadrature

For $v^H, w^H \in V^1(\Omega, \mathcal{T}_H)$, let us introduce the bilinear form $B_0 = B_{D,0} + B_{A,0}$ by

$$\begin{aligned} B_{D,0}(v^H, w^H) &= \int_\Omega a^0(x)\nabla v^H(x)\cdot\nabla w^H(x)dx - \int_\Gamma \{a^0(s)\nabla v^H(s)\}\cdot[\![w^H]\!]ds + \int_\Gamma \mu_S[\![v^H]\!]\cdot[\![w^H]\!]ds, \\ B_{A,0}(v^H, w^H) &= \int_\Omega b^0(x)\cdot\nabla v^H(x)\, w^H(x)dx - \int_\Gamma b^0(s)\cdot[\![v^H]\!]w^H_{i_0(s)}ds, \end{aligned} \tag{58}$$

with the penalty weighting function μ_S on an edge $e \in \mathcal{E}$ given by $\mu_S|_e = \alpha\big\|\{a^0(s)\}\big\|_{\mathcal{F}} H_e^{-1}$, where the penalization parameter $\alpha > 1$ is a positive parameter independent of the local mesh size and the data a^0, and the index $i_0(s)$ is discussed in Remark 3.2. We define $u^{0,H}$ as the solution of the variational problem: find $u^{0,H} \in V^1(\Omega, \mathcal{T}_H)$ such that

$$B_0(u^{0,H}, v^H) = \int_\Omega f\, v^H dx \qquad \forall\, v^H \in V^1(\Omega, \mathcal{T}_H). \tag{59}$$

Compared to the bilinear form $\tilde{B}_0$, defined in (48), the integrals are evaluated exactly in B_0, i.e., no numerical quadrature is used. Thus, the method given by (59) is free of any non-consistent perturbations and the Galerkin orthogonality holds

$$B_0(u^0 - u^{0,H}, v^H) = 0 \qquad \forall\, v^H \in V^1(\Omega, \mathcal{T}_H). \tag{60}$$

The stability of the method can be shown following the proof of Theorem 4.3 by setting $r_{vc,A} = 0$ (cf. Remark 5.10). Moreover, the a priori error estimate can be derived analogously to [18, Theorem 5.1], i.e., if $u^0 \in H^2(\Omega)$ then

$$|||u^0 - u^{0,H}||| \leq C(a_\infty^{1/2} H + b_\infty^{1/2} H^{3/2}) \big|u^0\big|_{H^2(\Omega)}. \tag{61}$$

A.2 DG-FEM with numerical quadrature

In this section, we study the single scale DG-FEM based on the bilinear form $\tilde{B}_0$ given by (48).
Stability. The proof of the inf-sup condition for $\tilde{B}_0$ follows the proof of Theorem 4.3 by replacing a_K^0 and b_K^0 by $a^0(x_K)$ and $b^0(x_K)$, respectively, leading to

$$\tilde{r}_{vc,A} = \frac{1}{b_\infty} \sup_{\substack{K \in \mathcal{T}_H \\ x \in K}} \big|b^0(x_K) - b^0(x)\big| + \frac{1}{b_\infty} \sup_{\substack{e \in \mathcal{E} \\ x \in e}} \big|\{b^0(x_K)\} - b^0(x)\big|,$$

This yields the same conditions `A`, `B`, `C` and `D` as in the proof of Theorem 4.3 with $r_{vc,A}$ replaced by $\tilde{r}_{vc,A}$.
A priori error estimate. Having shown the inf-sup condition for $\tilde{B}_0$ we derive the a priori error estimate for the single scale DG-FEM based on numerical integration used as estimate for the macro error e_{mac}.

Theorem A.1. *Let $u^0 \in H^2(\Omega)$, $a^0 \in (W^{1,\infty}(\Omega))^{d\times d}$ and $b^0 \in (W^{2,\infty}(\Omega))^d$. Then the solution $\tilde{u}^{0,H}$ of the variational problem (49) satisfies the estimate*

$$|||u^0 - \tilde{u}^{0,H}||| \leq C\Big(a_\infty^{1/2} H + b_\infty^{1/2} H^{3/2} + b_\infty^{1/2} \min\Big\{\mathrm{Pe}^{1/2}\, H^2, H\Big\}\Big) \big\|u^0\big\|_{H^2(\Omega)},$$

where C is independent of H.

Proof. We combine the ideas of the proof of Theorem 4.6 and [18, Theorem 5.1]. We decompose the total error into two parts $|||u^0 - \tilde{u}^{0,H}||| \leq |||u^0 - P_H u^0||| + |||P_H u^0 - \tilde{u}^{0,H}|||$ using the L^2 projection $P_H u^0$. Then, using the inf-sup condition for $\tilde{B}_0$, with stability constant $\tilde{\alpha}_S$, and the consistency (60) leads to

$$\begin{aligned}
\tilde{\alpha}_S |||P_H u^0 - \tilde{u}^{0,H}||| &\leq \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{\tilde{B}_0(P_H u^0 - \tilde{u}^{0,H}, w^H)}{|||w^H|||} = \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{\tilde{B}_0(P_H u^0, w^H) - B_0(u^{0,H}, w^H)}{|||w^H|||} \\
&= \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{\tilde{B}_0(P_H u^0, w^H) - B_0(u^0, w^H)}{|||w^H|||}.
\end{aligned}$$

Thus, we get the error decomposition

$$\begin{aligned}
|||u^0 - \tilde{u}^{0,H}||| \leq |||u^0 - P_H u^0||| &+ \frac{1}{\tilde{\alpha}_S} \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{B_0(P_H u^0 - u^0, w^H)}{|||w^H|||} \\
&+ \frac{1}{\tilde{\alpha}_S} \sup_{w^H \in V^1(\Omega,\mathcal{T}_H)} \frac{\tilde{B}_0(P_H u^0, w^H) - B_0(P_H u^0, w^H)}{|||w^H|||},
\end{aligned}$$

where the first two terms are identical to the error terms arising in the proof of (61) and the third term quantifies the effect of the numerical integration. Due to the decompositions $B_0 = B_{D,0} + B_{A,0}$ and $\tilde{B}_0 = \tilde{B}_{D,0} + \tilde{B}_{A,0}$ given by (58) and (48), respectively, we first estimate the difference $\tilde{B}_{D,0}(P_H u^0, w^H) - B_{D,0}(P_H u^0, w^H)$. Following the ideas of Lemma 5.12 we have

$$\begin{aligned}
\Big|\tilde{B}_{D,0}(P_H u^0, w^H) - B_{D,0}(P_H u^0, w^H)\Big| &\leq C\big\|a^0\big\|_{W^{1,\infty}(\Omega)} H \Big(\big\|\nabla P_H u^0\big\|_{L^2(\Omega)}^2 + \big|P_H u^0\big|_{*,D}^2\Big)^{1/2} |||w^H|||_D \\
&\leq C a_\infty^{1/2} H \big\|u^0\big\|_{H^2(\Omega)} |||w^H|||_D.
\end{aligned} \tag{62}$$

Next, we need to estimate $\tilde{B}_{A,0}(P_H u^0, w^H) - B_{A,0}(P_H u^0, w^H)$. Following Lemma 5.7 we obtain

$$\begin{aligned}
&\left| \sum_{K\in\mathcal{T}_H} \int_K \big(b^0(x_K) - b^0(x)\big) \cdot \nabla P_H u^0 \, w^H dx \right| \\
&\qquad \le C \big\|b^0\big\|_{W^{1,\infty}(\Omega)} H \big\|\nabla P_H u^0\big\|_{L^2(\Omega)} \big\|w^H\big\|_{L^2(\Omega)} \le C b_\infty^{1/2} H \big\|u^0\big\|_{H^2(\Omega)} |\!|\!| w^H |\!|\!|_A. \qquad (63) \\
&\left| \int_\Gamma \{b^0(x_K)\} \cdot [\![P_H u^0]\!] w^H_{\tilde{\imath}_0} ds - \int_\Gamma b^0(s) \cdot [\![P_H u^0]\!] w^H_{i_0(s)} ds \right| \\
&\qquad \le C \alpha^{-1/2} \sup_{e\in\mathcal{E}, x\in e} \big|\{b^0(x_K)\} - b^0(x)\big| \big|P_H u^0 - u^0\big|_{*,D} \big\|w^H\big\|_{L^2(\Gamma)} \le C b_\infty^{1/2} H^2 \big|u^0\big|_{H^2(\Omega)} |\!|\!| w^H |\!|\!|_A, \qquad (64)
\end{aligned}$$

where we used $\big|P_H u^0\big|_{*,D} = \big|P_H u^0 - u^0\big|_{*,D}$, as $H^2(\Omega) \hookrightarrow C^0(\Omega)$ for $d \le 3$ and $u^0 = 0$ on $\partial\Omega$. If b^0 has the additional regularity $b^0 \in (W^{2,\infty}(\Omega))^d$, we can improve estimate (63) using [23, Theorem 4]

$$\begin{aligned}
\left| \sum_{K\in\mathcal{T}_H} \int_K \big(b^0(x_K) - b^0(x)\big) \cdot \nabla P_H u^0 \, w^H dx \right| &\le C \big\|b^0\big\|_{W^{2,\infty}(\Omega)} H^2 \big\|\nabla P_H u^0\big\|_{L^2(\Omega)} \big\|\nabla w^H\big\|_{L^2(\Omega)} \\
&\le C \frac{b_\infty}{a_\infty^{1/2}} H^2 \big\|u^0\big\|_{H^2(\Omega)} |\!|\!| w^H |\!|\!|_D. \qquad (65)
\end{aligned}$$

Finally, combining estimates (62), (63), (64) and (65) allows to estimate the effect of the quadrature

$$\Big|\tilde{B}_0(P_H u^0, w^H) - B_0(P_H u^0, w^H)\Big| \le C\Big(a_\infty^{1/2} H + b_\infty^{1/2} \min\Big\{\mathrm{Pe}^{1/2} H^2, H\Big\} + b_\infty^{1/2} H^2\Big) \big\|u^0\big\|_{H^2(\Omega)} |\!|\!| w^H |\!|\!|. \; \square$$