Comput. Methods Appl. Math. 2019; 19(4): 723–748

Research Article

Matteo Cicuttin, Alexandre Ern and Simon Lemaire*

A Hybrid High-Order Method for Highly Oscillatory Elliptic Problems

https://doi.org/10.1515/cmam-2018-0013
Received February 15, 2018; accepted June 1, 2018

Abstract: We devise a Hybrid High-Order (HHO) method for highly oscillatory elliptic problems that is capable of handling general meshes. The method hinges on discrete unknowns that are polynomials attached to the faces and cells of a coarse mesh; those attached to the cells can be eliminated locally using static condensation. The main building ingredient is a reconstruction operator, local to each coarse cell, that maps onto a fine-scale space spanned by oscillatory basis functions. The present HHO method generalizes the ideas of some existing multiscale approaches, while providing the first complete analysis on general meshes. It also improves on those methods, taking advantage of the flexibility granted by the HHO framework. The method handles arbitrary orders of approximation $k \geq 0$. For face unknowns that are polynomials of degree k, we devise two versions of the method, depending on the polynomial degree $(k-1)$ or k of the cell unknowns. We prove, in the case of periodic coefficients, an energy-error estimate of the form $(\varepsilon^{\frac{1}{2}} + H^{k+1} + (\frac{\varepsilon}{H})^{\frac{1}{2}})$, and we illustrate our theoretical findings on some test-cases.

Keywords: General Meshes, HHO Methods, Multiscale Methods, Highly Oscillatory Problems

MSC 2010: 65N30, 65N08, 76R50

1 Introduction

Over the last few years, many advances have been accomplished in the design of arbitrary-order polytopal discretization methods. Such methods are capable of handling meshes with polytopal cells, and possibly including hanging nodes. The use of polytopal meshes can be motivated by the increased flexibility, when meshing complex geometries, or when using agglomeration techniques for mesh coarsening (see, e.g., [7]). Classical examples of polytopal methods are the (polytopal) Finite Element Method (FEM) [44, 46], which typically uses non-polynomial basis functions to enforce continuity, and non-conforming methods such as the Discontinuous Galerkin (DG) [5, 10, 16] and the Hybridizable Discontinuous Galerkin (HDG) [15] methods. We also mention the Weak Galerkin (WG) [47] method (see [13] for its links to HDG).

More recently, new paradigms have emerged. One salient example is the Virtual Element Method (VEM) [9], which is formulated in terms of virtual (i.e., non-computed) conforming functions. The key idea is that the virtual space contains those polynomial functions leading to optimal approximation properties, whereas the remaining functions need not be computed (only their degrees of freedom need to be) provided some suitable local stabilization is introduced. The degrees of freedom in the VEM are attached to the mesh vertices, and, as the order of the approximation is increased, also to the mesh edges, faces, and cells. Another

Matteo Cicuttin, Alexandre Ern, Université Paris-Est, CERMICS (ENPC), 6–8 avenue Blaise Pascal, 77455 Marne-la-Vallée Cedex 2; and Inria Paris, 2 rue Simone Iff, 75012 Paris, France, e-mail: matteo.cicuttin@enpc.fr, alexandre.ern@enpc.fr
***Corresponding author: Simon Lemaire,** École Polytechnique Fédérale de Lausanne, FSB-MATH-ANMC, Station 8, 1015 Lausanne, Switzerland; and Inria Lille – Nord Europe, 40 avenue Halley, 59650 Villeneuve d'Ascq, France, e-mail: simon.lemaire@inria.fr

recent polytopal method is the Hybrid High-Order (HHO) method, which has been introduced for locking-free linear elasticity in [17], and for diffusion in [19]. The HHO method has been formulated originally as a non-conforming method, using polynomial unknowns attached to the mesh faces and cells. The HHO method has been bridged in [14] both to HDG (by identifying a suitable numerical flux trace), and to the non-conforming VEM considered in [6] (by identifying an isomorphism between the HHO degrees of freedom and a local virtual finite-dimensional space, which again contains those polynomial functions leading to optimal approximation properties). The focus here is on HHO methods. HHO methods offer several assets, including a dimension-independent construction, local conservativity, and attractive computational costs, especially in 3D. Indeed, the HHO stencil is more compact than for methods involving degrees of freedom attached to the mesh vertices, and static condensation allows one to eliminate cell degrees of freedom, leading to a global problem expressed in terms of face degrees of freedom only, whose number grows quadratically with the polynomial order, whereas the growth of globally coupled degrees of freedom is typically cubic for DG methods.

In this work, we are interested in elliptic problems featuring heterogeneous/anisotropic coefficients that are highly oscillatory. The case of slowly varying coefficients has already been treated in [18, 20], where error estimates tracking the dependency of the approximation with respect to the local heterogeneity/anisotropy ratios have been derived. Let Ω be an open, bounded, connected polytopal subset of $\mathbb{R}^d$, $d \in \{2,3\}$, and $\varepsilon > 0$, supposedly much smaller than the diameter of the domain Ω, encode the highly oscillatory nature of the coefficients. We consider the model problem

$$\begin{cases} -\operatorname{div}(\mathbb{A}_\varepsilon \nabla u_\varepsilon) = f & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where $f \in L^2(\Omega)$ is non-oscillatory, and $\mathbb{A}_\varepsilon$ is an oscillatory, uniformly elliptic and bounded matrix-valued field on Ω. It is well known that the H^{k+2}-norm of the solution u_ε to problem (1.1) scales as $\varepsilon^{-(k+1)}$, meaning that monoscale methods (including the monoscale HHO method of order $k \geq 0$ of [18, 20]) provide an energy-norm decay of the error of order $(\frac{h}{\varepsilon})^{k+1}$. To be accurate, such methods must hence rely on a mesh resolving the fine scale, i.e., with size $h \ll \varepsilon$. Since ε is supposedly much smaller than the diameter of Ω, an accurate approximation necessarily implies an overwhelming number of degrees of freedom. In a multi-query context, where the solution is needed for a large number of right-hand sides (e.g., an optimization loop, with f as a control and (1.1) as a distributed constraint), a monoscale solve is hence unaffordable. In that context, multiscale methods may be preferred. Multiscale methods aim at resolving the fine scale in an offline step, reducing the online step to the solution of a system of small size, based on oscillatory basis functions computed in the offline step, on a coarse mesh with size $H \gg \varepsilon$. In a single-query context, multiscale methods are also interesting since they allow one to organize computations in a more efficient way.

Multiscale approximation methods on classical element shapes (such as simplices or quadrangles/hexahedra) have been analyzed extensively in the literature. Examples include, e.g., the multiscale Finite Element Method (msFEM) [23, 34, 35] (with energy-error bound of the form $(\varepsilon^{\frac{1}{2}} + H + (\frac{\varepsilon}{H})^{\frac{1}{2}})$ in the periodic case), its variant using oversampling [24, 34] (with improved error bound of the form $(\varepsilon^{\frac{1}{2}} + H + \frac{\varepsilon}{H})$ in the periodic case), or the Petrov–Galerkin variant of the msFEM using oversampling [36]. Let us also mention [3] (see also [33]), which is an extension to arbitrary orders of approximation of the classical msFEM (with error bound of the form $(\varepsilon^{\frac{1}{2}} + H^k + (\frac{\varepsilon}{H})^{\frac{1}{2}})$ in the periodic case using H^1-conforming finite elements of degree $k \geq 1$). These methods all rely on the assumption that a conforming finite element basis is available for the (coarse) mesh under consideration. Recent research directions essentially focus on the approximation of problems that do not assume scale separation, and on reducing and possibly eliminating the cell resonance error. One can cite, e.g., the Generalized msFEM (GmsFEM) [22], or the Local Orthogonal Decomposition (LOD) approach [32, 41]. We also mention that other paradigms exist to approximate oscillatory problems, like the Heterogeneous Multiscale Method (HMM) [1, 21].

On general polytopal meshes, the literature on multiscale methods is more scarce. For constructions in the spirit of the msFEM, one can cite the msFEM à la Crouzeix–Raviart of [39, 40], the so-called Multiscale Hybrid-Mixed (MHM) [4, 43] approach, and the (polynomial-based) method of [26] in the HDG context. Each one of these methods has its proper design, but they all share the same construction principles: they are based, more or less directly, on oscillatory basis functions that solve local Neumann problems with polynomial boundary data, and result in global systems (posed on the coarse mesh) that can be expressed in terms of face unknowns only. In the following, we will thus refer to those methods as skeletal-msFEM. The MHM approach actually presents a small difference with respect to the two other approaches since it is based on a hybridized primal formulation, which leads to consider flux-type unknowns at interfaces instead of potential-type unknowns; as a consequence, and in order to impose the compatibility constraint, one needs to solve a saddle-point global problem, whereas for the two other approaches, one ends up with a coercive problem. For the msFEM à la Crouzeix–Raviart, an error bound of the form $(\varepsilon^{\frac{1}{2}} + H + (\frac{\varepsilon}{H})^{\frac{1}{2}})$ is proved in [39] in the periodic case. However, the analysis is led under the assumption that there exists a finite number of reference elements in the mesh sequence. For the MHM approach, which is designed in the same spirit, the same type of upper bound for the error is expected. Yet, in [43], the authors claim that their method is able to get rid of the resonance error (without oversampling); we clarify this issue in Remark A.3 below. For the HDG-like method, the analysis that is provided in [26] is sharp only in the regime $H \ll \varepsilon$. As a consequence, there is, to date, no complete polytopal analysis available in the literature for skeletal-msFEM. Moreover, we observe that in the three methods, the discretization of the (non-oscillatory) right-hand side is realized in a somewhat suboptimal way, which can become a limiting issue in a multi-query context. In the msFEM à la Crouzeix–Raviart, the discretization is realized through a projection of the loading term onto the space spanned by the oscillatory basis functions. In the MHM and HDG-like approaches, the whole (local) space H^1 is considered. In all cases, the approximation of the right-hand side does not take advantage of the fact that the latter is non-oscillatory. Let us mention, as another construction in the spirit of the msFEM, the work [38], which exploits in the DG context the ideas introduced in [3]. The drawback, which is inherent to DG methods, is the large size of the online systems. For constructions in the spirit of the GmsFEM, let us mention in the HDG context the contributions [11, 25] (that are based on [26]), and the work [42] in the WG context.

In this work, we devise a multiscale HHO (msHHO) method, which can be seen as a generalization (in particular to arbitrary orders of approximation) of the msFEM à la Crouzeix–Raviart of [39, 40]. Our contribution is twofold. First, we provide an analysis (in the periodic setting) of the method that is valid on general polytopal mesh sequences (in particular, we do not postulate the existence of reference elements); in that respect, this work presents the first complete polytopal analysis of a skeletal-msFEM. Note that considering general element shapes in the periodic setting is clearly not a good strategy (cf., e.g., [31]); however, this setting is not our final target. Second, taking advantage of the flexibility offered by the HHO framework, we improve on the existing methods. We introduce (polynomial) cell unknowns, that we use for the integration of the right-hand side (cf. Remarks 5.8 and 5.16 below). The non-oscillatory loading is hence discretized through a coarse-scale polynomial projection, while the size of the online system remains unchanged since the cell unknowns are locally eliminated in the offline step. Two versions of the msHHO method are proposed herein, both employing polynomials of arbitrary order $k \geq 0$ for the face unknowns. For the mixed-order msHHO method, the cell unknowns are polynomials of order $(k-1)$ (if $k \geq 1$), whereas they are polynomials of order $k \geq 0$ for the equal-order msHHO method. The mixed-order msHHO method does not require stabilization, whereas a simple stabilization (which avoids computing additional oscillatory basis functions) is introduced in the equal-order case. We prove for both methods an energy-error estimate of the form $(\varepsilon^{\frac{1}{2}} + H^{k+1} + (\frac{\varepsilon}{H})^{\frac{1}{2}}) =: g_k(H)$ in the periodic case. The analysis of the msHHO method differs from that of the monoscale HHO method since the local fine-scale space does not contain polynomial functions up to order $(k+1)$; in this respect, our key approximation result is Lemma 4.5 below. With respect to [39], we also simplify the analysis and weaken the regularity assumptions (cf. Remark 4.6 below). Our analysis finally sheds new light on the relationship between the non-computed functions of the local virtual space and the associated local stabilization. To motivate the design and use of a high-order method, we note, as it was already pointed out in [3], that the upper bound $g_k(H)$ is minimal for $H_k = (\varepsilon^{\frac{1}{2}}/2(k+1))^{2/(2k+3)}$, hence as $k \geq 0$ increases, H_k increases whereas $g_k(H_k)$ decreases. The msHHO method we devise is meant to be a first step in the design of an accurate and computationally effective multiscale approach on general meshes. The next step will be to address the resonance phenomenon and the more realistic setting of no scale separation.

The article is organized as follows. In Sections 2 and 3 we introduce, respectively, the continuous and discrete settings. In Section 4, we introduce the fine-scale approximation space, exhibiting its (oscillatory) basis functions and studying, locally, its approximation properties. In Section 5, we introduce the two versions of the msHHO method, analyze their stability, and derive energy-error estimates. In Section 6, we present some numerical illustrations in the periodic and locally periodic settings. Finally, in Appendix A we collect some useful estimates on the first-order two-scale expansion.

2 Continuous Setting

From now on, and in order to lead the analysis, we assume that the diffusion matrix $\mathbb{A}_\varepsilon$ satisfies $\mathbb{A}_\varepsilon(\cdot) = \mathbb{A}(\cdot/\varepsilon)$ in Ω, where $\mathbb{A}$ is a symmetric and $\mathbb{Z}^d$-periodic matrix field on $\mathbb{R}^d$. Letting $Q := (0,1)^d$, we define, for $1 \le p \le +\infty$ and $m \in \mathbb{N}^\star$, the following periodic spaces:

$$L^p_{\mathrm{per}}(Q) := \{v \in L^p_{\mathrm{loc}}(\mathbb{R}^d) : v \text{ is } \mathbb{Z}^d\text{-periodic}\},$$
$$W^{m,p}_{\mathrm{per}}(Q) := \{v \in W^{m,p}_{\mathrm{loc}}(\mathbb{R}^d) : v \text{ is } \mathbb{Z}^d\text{-periodic}\},$$

with the classical conventions that $W^{m,2}_{\mathrm{per}}(Q)$ is denoted $H^m_{\mathrm{per}}(Q)$ and that the subscript "loc" can be omitted for $p = +\infty$. Letting $\mathcal{S}_d(\mathbb{R})$ denote the set of real-valued $d \times d$ symmetric matrices, we also define, for real numbers $0 < a \le b$,

$$\mathcal{S}_a^b := \{\mathbb{M} \in \mathcal{S}_d(\mathbb{R}) : a|\boldsymbol{\xi}|^2 \le \mathbb{M}\boldsymbol{\xi}\cdot\boldsymbol{\xi} \le b|\boldsymbol{\xi}|^2 \text{ for all } \boldsymbol{\xi} \in \mathbb{R}^d\}.$$

We assume that there exist real numbers $0 < \alpha \le \beta$ such that

$$\mathbb{A}(\cdot) \in \mathcal{S}_\alpha^\beta \quad \text{a.e. in } \mathbb{R}^d. \tag{2.1}$$

Assumption (2.1) ensures that $\mathbb{A}_\varepsilon \in L^\infty(\Omega; \mathbb{R}^{d\times d})$ is such that $\mathbb{A}_\varepsilon(\cdot) \in \mathcal{S}_\alpha^\beta$ a.e. in Ω for any $\varepsilon > 0$, and hence guarantees the existence and uniqueness of the solution to (1.1) in $H^1_0(\Omega)$ for any $\varepsilon > 0$. More importantly, assumption (2.1) ensures that the (whole) family $(\mathbb{A}_\varepsilon)_{\varepsilon>0}$ G-converges [2, Section 1.3.2] to some constant symmetric matrix $\mathbb{A}_0 \in \mathcal{S}_\alpha^\beta$. Henceforth, we denote $\rho := \frac{\beta}{\alpha} \ge 1$ the (global) heterogeneity/anisotropy ratio of both $(\mathbb{A}_\varepsilon)_{\varepsilon>0}$ and $\mathbb{A}_0$. Letting $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_d)$ denote the canonical basis of $\mathbb{R}^d$, the expression of $\mathbb{A}_0$ is known to read, for integers $1 \le i, j \le d$,

$$[\mathbb{A}_0]_{ij} = \int_Q \mathbb{A}(\boldsymbol{e}_j + \nabla\mu_j)\cdot(\boldsymbol{e}_i + \nabla\mu_i) = \int_Q \mathbb{A}(\boldsymbol{e}_j + \nabla\mu_j)\cdot\boldsymbol{e}_i, \tag{2.2}$$

where, for any integer $1 \le l \le d$, the so-called corrector $\mu_l \in H^1_{\mathrm{per}}(Q)$ is the solution with zero mean-value on Q to the problem

$$\begin{cases} -\mathrm{div}(\mathbb{A}(\nabla\mu_l + \boldsymbol{e}_l)) = 0 & \text{in } \mathbb{R}^d, \\ \mu_l \text{ is } \mathbb{Z}^d\text{-periodic}. \end{cases} \tag{2.3}$$

For further use, we also define the linear operator $\mathcal{R}_\varepsilon : L^p_{\mathrm{per}}(Q) \to L^p(\Omega)$, $1 \le p \le +\infty$, such that, for any function $\chi \in L^p_{\mathrm{per}}(Q)$, $\mathcal{R}_\varepsilon(\chi) \in L^p(\Omega)$ satisfies $\mathcal{R}_\varepsilon(\chi)(\cdot) = \chi(\cdot/\varepsilon)$ in Ω. In particular, for any integers $1 \le i, j \le d$, we have $[\mathbb{A}_\varepsilon]_{ij} = \mathcal{R}_\varepsilon(\mathbb{A}_{ij})$. A useful property of $\mathcal{R}_\varepsilon$ is the relation $\partial_l(\mathcal{R}_\varepsilon(\chi)) = \frac{1}{\varepsilon}\mathcal{R}_\varepsilon(\partial_l\chi)$, valid for any function $\chi \in W^{1,p}_{\mathrm{per}}(Q)$ and any integer $1 \le l \le d$.

The homogenized problem reads

$$\begin{cases} -\mathrm{div}(\mathbb{A}_0\nabla u_0) = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega. \end{cases} \tag{2.4}$$

We introduce the so-called first-order two-scale expansion

$$\mathcal{L}^1_\varepsilon(u_0) := u_0 + \varepsilon\sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l)\partial_l u_0. \tag{2.5}$$

Note that $(u_\varepsilon - \mathcal{L}^1_\varepsilon(u_0))$ does not a priori vanish on the boundary of Ω.

3 Discrete Setting

We denote by $\mathcal{H} \subset \mathbb{R}_+^\star$ a countable set of meshsizes having 0 as its unique accumulation point, and we consider mesh sequences of the form $(\mathcal{T}_H)_{H\in\mathcal{H}}$. For any meshsize $H \in \mathcal{H}$, a *mesh* $\mathcal{T}_H$ is a finite collection of nonempty disjoint open polytopes (polygons/polyhedra) T, called *elements* or *cells*, such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}_H} \overline{T}$ and $H = \max_{T\in\mathcal{T}_H} H_T$, H_T standing for the diameter of the cell T. The mesh cells being polytopal, their boundary is composed of a finite union of portions of affine hyperplanes in $\mathbb{R}^d$ called facets (each facet has positive $(d-1)$-dimensional measure). A closed subset F of $\overline{\Omega}$ is called a *face* if either

(i) there exist $T_1, T_2 \in \mathcal{T}_H$ such that $F = \partial T_1 \cap \partial T_2 \cap Z$, where Z is an affine hyperplane supporting a facet of both T_1 and T_2 (and F is termed *interface*), or

(ii) there exists $T \in \mathcal{T}_H$ such that $F = \partial T \cap \partial\Omega \cap Z$, where Z is an affine hyperplane supporting a facet of both T and Ω (and F is termed *boundary face*).

Interfaces are collected in the set $\mathcal{F}_H^{\mathrm{i}}$, boundary faces in $\mathcal{F}_H^{\mathrm{b}}$, and we let $\mathcal{F}_H := \mathcal{F}_H^{\mathrm{i}} \cup \mathcal{F}_H^{\mathrm{b}}$. The diameter of a face $F \in \mathcal{F}_H$ is denoted H_F. For all $T \in \mathcal{T}_H$, we define

$$\mathcal{F}_T := \{F \in \mathcal{F}_H : F \subset \partial T\}$$

the set of faces lying on the boundary of T; note that the faces in $\mathcal{F}_T$ compose the boundary of T. For any $T \in \mathcal{T}_H$, we denote by $\boldsymbol{n}_{\partial T}$ the unit normal vector to ∂T pointing outward T, and for any $F \in \mathcal{F}_T$, we let $\boldsymbol{n}_{T,F} := \boldsymbol{n}_{\partial T|F}$ (by definition, $\boldsymbol{n}_{T,F}$ is a constant vector on F).

We adopt the following notion of admissible mesh sequence; cf. [16, Section 1.4] and [20, Definition 2.1].

Definition 3.1 (Admissible Mesh Sequence). The mesh sequence $(\mathcal{T}_H)_{H\in\mathcal{H}}$ is *admissible* if, for all $H \in \mathcal{H}$, $\mathcal{T}_H$ admits a matching simplicial sub-mesh $\mathfrak{T}_H$ (meaning that the cells in $\mathfrak{T}_H$ are sub-cells of the cells in $\mathcal{T}_H$ and that the faces of these sub-cells belonging to the skeleton of $\mathcal{T}_H$ are sub-faces of the faces in $\mathcal{F}_H$), and there exists a real number $\gamma > 0$, called *mesh regularity parameter*, such that, for all $H \in \mathcal{H}$, the following holds:

(i) For all simplex $S \in \mathfrak{T}_H$ of diameter H_S and inradius R_S, $\gamma H_S \le R_S$.

(ii) For all $T \in \mathcal{T}_H$, and all $S \in \mathfrak{T}_T := \{S \in \mathfrak{T}_H : S \subseteq T\}$, $\gamma H_T \le H_S$.

Two classical consequences of Definition 3.1 are that, for any mesh $\mathcal{T}_H$ belonging to an admissible mesh sequence,

(i) the quantity $\operatorname{card}(\mathcal{F}_T)$ is bounded independently of the diameter H_T for all $T \in \mathcal{T}_H$ (see [16, Lemma 1.41]),

(ii) mesh faces have a comparable diameter to the diameter of the cells to which they belong (see [16, Lemma 1.42]).

For any $q \in \mathbb{N}$, and any integer $1 \le l \le d$, we denote by $\mathbb{P}_l^q$ the linear space spanned by l-variate polynomial functions of total degree less than or equal to q. We let

$$\mathrm{N}_l^q := \dim(\mathbb{P}_l^q) = \binom{q+l}{q}.$$

Let a mesh $\mathcal{T}_H$ be given. For any $T \in \mathcal{T}_H$, $\mathbb{P}_d^q(T)$ is composed of the restriction to T of polynomials in $\mathbb{P}_d^q$, and for any $F \in \mathcal{F}_H$, $\mathbb{P}_{d-1}^q(F)$ is composed of the restriction to F of polynomials in $\mathbb{P}_d^q$ (this space can also be described as the restriction to F of polynomials in $\mathbb{P}_{d-1}^q \circ \Theta^{-1}$, where Θ is any affine bijective mapping from $\mathbb{R}^{d-1}$ to the affine hyperplane supporting F). We also introduce, for any $T \in \mathcal{T}_H$, the following broken polynomial space:

$$\mathbb{P}_{d-1}^q(\mathcal{F}_T) := \{v \in L^2(\partial T) : v_{|F} \in \mathbb{P}_{d-1}^q(F) \text{ for all } F \in \mathcal{F}_T\}.$$

The term “broken” refers to the fact that no continuity is required between adjacent faces for functions in $\mathbb{P}_{d-1}^q(\mathcal{F}_T)$. For any $T \in \mathcal{T}_H$, we denote by $(\Phi_T^{q,i})_{1\le i\le \mathrm{N}_d^q}$ a set of basis functions of the space $\mathbb{P}_d^q(T)$, and for any $F \in \mathcal{F}_H$, we denote by $(\Phi_F^{q,j})_{1\le j\le \mathrm{N}_{d-1}^q}$ a set of basis functions of the space $\mathbb{P}_{d-1}^q(F)$. We define, for any $T \in \mathcal{T}_H$ and $F \in \mathcal{F}_H$, Π_T^q and Π_F^q as the L^2-orthogonal projectors onto the spaces $\mathbb{P}_d^q(T)$ and $\mathbb{P}_{d-1}^q(F)$, respectively. Whenever no confusion can arise, we write, for all $T \in \mathcal{T}_H$, all $F \in \mathcal{F}_T$, and all $v \in H^1(T)$, $\Pi_F^q(v)$ instead of $\Pi_F^q(v_{|F})$.

We conclude this section by recalling some classical results, that are valid for any mesh $\mathcal{T}_H$ belonging to an admissible mesh sequence in the sense of Definition 3.1. For any $T \in \mathcal{T}_H$ and $F \in \mathcal{F}_T$, the trace inequalities

$$\|v\|_{L^2(F)} \le c_{\mathrm{tr,d}} H_F^{-\frac{1}{2}} \|v\|_{L^2(T)} \quad \text{for all } v \in \mathbb{P}_d^q(T), \tag{3.1}$$

$$\|v\|_{L^2(F)} \le c_{\mathrm{tr,c}} (H_T^{-1} \|v\|_{L^2(T)}^2 + H_T \|\nabla v\|_{L^2(T)^d}^2)^{\frac{1}{2}} \quad \text{for all } v \in H^1(T), \tag{3.2}$$

hold [16, Lemmas 1.46 and 1.49], as well as the local Poincaré inequality

$$\|v\|_{L^2(T)} \le c_{\mathrm{P}} H_T \|\nabla v\|_{L^2(T)^d} \quad \text{for all } v \in H^1(T) \text{ such that } \int_T v = 0, \tag{3.3}$$

where $c_{\mathrm{P}} = \pi^{-1}$ for convex elements [8]; estimates in the non-convex case can be found, e.g., in [45]. Finally, proceeding as in [27, Lemma 5.6], one can prove using the above trace and Poincaré inequalities that

$$|v - \Pi_T^q(v)|_{H^m(T)} + H_T^{\frac{1}{2}} |v - \Pi_T^q(v)|_{H^m(F)} \le c_{\mathrm{app}} H_T^{s-m} |v|_{H^s(T)} \quad \forall v \in H^s(T), \tag{3.4}$$

for integers $1 \le s \le q+1$ and $0 \le m \le (s-1)$. All of the above constants are independent of the meshsize and can only depend on the underlying polynomial degree q, the space dimension d, and the mesh regularity parameter γ.

Henceforth, we use the symbol c to denote a generic positive constant, whose value can change at each occurrence, provided it is independent of the micro-scale ε, any meshsize H_T or H, and the homogenized solution u_0. We also track the direct dependency of the error bounds on the parameters α, β characterizing the spectrum of the diffusion matrix. The value of the generic constant c can depend on the space dimension d, the underlying polynomial degree, the mesh regularity parameter γ, and on some higher-order norms of the rescaling $\frac{\mathbb{A}}{\beta}$ of the diffusion matrix or the correctors μ_l that will be made clear from the context.

4 Fine-Scale Approximation Space

Let $k \in \mathbb{N}$ and let $\mathcal{T}_H$ be a member of an admissible mesh sequence in the sense of Definition 3.1. In this section, we introduce the fine-scale approximation space on which we will base our multiscale HHO method. We first construct in Section 4.1 a set of cell-based and face-based basis functions, then we provide in Section 4.2 a local characterization of the underlying space, finally we study its approximation properties in Section 4.3.

4.1 Oscillatory Basis Functions

The oscillatory basis functions consist of cell- and face-based basis functions.

4.1.1 Cell-Based Basis Functions

Let $T \in \mathcal{T}_H$. If $k = 0$, we do not define cell-based basis functions. Assume now that $k \ge 1$. For all $1 \le i \le \mathrm{N}_d^{k-1}$, we consider the problem

$$\inf \left\{ \int_T \left[\frac{1}{2} \mathbb{A}_\varepsilon \nabla\varphi \cdot \nabla\varphi - \Phi_T^{k-1,i} \varphi \right] : \varphi \in H^1(T),\ \Pi_F^k(\varphi) = 0 \text{ for all } F \in \mathcal{F}_T \right\}. \tag{4.1}$$

Problem (4.1) admits a unique minimizer. This minimizer, that we will denote $\varphi_{\varepsilon,T}^{k+1,i} \in H^1(T)$, can be proved to solve, for real numbers $(\lambda_{F,j}^T)_{F \in \mathcal{F}_T,\, 1 \le j \le \mathrm{N}_{d-1}^k}$ satisfying the compatibility condition

$$\sum_{F \in \mathcal{F}_T} \int_F \sum_{j=1}^{\mathrm{N}_{d-1}^k} \lambda_{F,j}^T \Phi_F^{k,j} = -\int_T \Phi_T^{k-1,i},$$

the constrained Neumann problem

$$\begin{cases} -\mathrm{div}(\mathbb{A}_\varepsilon \nabla \varphi_{\varepsilon,T}^{k+1,i}) = \Phi_T^{k-1,i} & \text{in } T, \\ \mathbb{A}_\varepsilon \nabla \varphi_{\varepsilon,T}^{k+1,i} \cdot \boldsymbol{n}_{T,F} = \sum_{j=1}^{\mathrm{N}_{d-1}^k} \lambda_{F,j}^T \Phi_F^{k,j} & \text{on all } F \in \mathcal{F}_T, \\ \Pi_F^k(\varphi_{\varepsilon,T}^{k+1,i}) = 0 & \text{on all } F \in \mathcal{F}_T. \end{cases} \tag{4.2}$$

The superscript $k+1$ is meant to remind us that the functions $\varphi_{\varepsilon,T}^{k+1,i}$ are used to generate a linear space which has the same approximation capacity as the polynomial space of order at most $k+1$, as will be shown in Section 4.3.

Remark 4.1 (Practical Computation). To compute the functions $\varphi_{\varepsilon,T}^{k+1,i}$ for all $1 \le i \le \mathrm{N}_d^{k-1}$, one considers in practice a (shape-regular) matching simplicial mesh $\mathcal{T}_h^T$ of the cell T, with size h smaller than ε. Then one can solve problem (4.2) approximately by using a classical (equal-order) monoscale HHO method (or any other monoscale approximation method). For the implementation of the monoscale HHO method, we refer to [12]. One can either consider a weak formulation in $\{\varphi \in H^1(T) : \Pi_F^k(\varphi) = 0 \text{ for all } F \in \mathcal{F}_T\}$, which leads to a coercive problem, or a weak formulation in $H^1(T)$, which leads to a saddle-point system with Lagrange multipliers. Equivalent considerations apply below to the computation of the face-based basis functions. Note that the error estimates we provide in this work for our approach do not take into account the local approximations of size h and assume that (4.2) and (4.4) below are solved exactly.

4.1.2 Face-Based Basis Functions

Let $T \in \mathcal{T}_H$. For all $F \in \mathcal{F}_T$ and all $1 \le j \le \mathrm{N}_{d-1}^k$, we consider the problem

$$\inf\left\{ \int_T \left[\frac{1}{2}\mathbb{A}_\varepsilon \nabla\varphi \cdot \nabla\varphi\right] : \varphi \in H^1(T),\ \Pi_F^k(\varphi) = \Phi_F^{k,j},\ \Pi_\sigma^k(\varphi) = 0 \text{ for all } \sigma \in \mathcal{F}_T \setminus \{F\} \right\}. \tag{4.3}$$

Problem (4.3) admits a unique minimizer. This minimizer, that we will denote $\varphi_{\varepsilon,T,F}^{k+1,j} \in H^1(T)$, can be proved to solve, for real numbers $(\lambda_{\sigma,q}^{T,F})_{\sigma\in\mathcal{F}_T,\,1\le q\le \mathrm{N}_{d-1}^k}$ satisfying the compatibility condition

$$\sum_{\sigma\in\mathcal{F}_T} \int_\sigma \sum_{q=1}^{\mathrm{N}_{d-1}^k} \lambda_{\sigma,q}^{T,F} \Phi_\sigma^{k,q} = 0,$$

the constrained Neumann problem

$$\begin{cases} -\mathrm{div}(\mathbb{A}_\varepsilon \nabla \varphi_{\varepsilon,T,F}^{k+1,j}) = 0 & \text{in } T, \\ \mathbb{A}_\varepsilon \nabla \varphi_{\varepsilon,T,F}^{k+1,j} \cdot \boldsymbol{n}_{T,\sigma} = \sum_{q=1}^{\mathrm{N}_{d-1}^k} \lambda_{\sigma,q}^{T,F} \Phi_\sigma^{k,q} & \text{on all } \sigma \in \mathcal{F}_T, \\ \Pi_F^k(\varphi_{\varepsilon,T,F}^{k+1,j}) = \Phi_F^{k,j} & \text{on } F, \\ \Pi_\sigma^k(\varphi_{\varepsilon,T,F}^{k+1,j}) = 0 & \text{on all } \sigma \in \mathcal{F}_T \setminus \{F\}. \end{cases} \tag{4.4}$$

4.2 Discrete Space

We introduce, for any $T \in \mathcal{T}_H$, the space

$$V_{\varepsilon,T}^{k+1} := \{v_\varepsilon \in H^1(T) : \mathrm{div}(\mathbb{A}_\varepsilon \nabla v_\varepsilon) \in \mathbb{P}_d^{k-1}(T),\ \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}_{d-1}^k(\mathcal{F}_T)\}, \tag{4.5}$$

with the convention that $\mathbb{P}_d^{-1}(T) := \{0\}$. We recall that the condition $\mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}_{d-1}^k(\mathcal{F}_T)$ is equivalent to $\mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{T,F} \in \mathbb{P}_{d-1}^k(F)$ for all $F \in \mathcal{F}_T$. Proceeding as in [14, Section 2.4], it can easily be shown that the dimension of $V_{\varepsilon,T}^{k+1}$ is $(\mathrm{N}_d^{k-1} + \mathrm{card}(\mathcal{F}_T) \times \mathrm{N}_{d-1}^k)$ (or $\mathrm{card}(\mathcal{F}_T)$ if $k = 0$).

Proposition 4.2 (Characterization of $V_{\varepsilon,T}^{k+1}$). *For any $T \in \mathcal{T}_H$, the family $\{(\varphi_{\varepsilon,T}^{k+1,i})_{1\le i\le \mathrm{N}_d^{k-1}}, (\varphi_{\varepsilon,T,F}^{k+1,j})_{F\in\mathcal{F}_T,\,1\le j\le \mathrm{N}_{d-1}^k}\}$ forms a basis for the space $V_{\varepsilon,T}^{k+1}$.*

Proof. To establish the result, we only need to prove that

$$V_{\varepsilon,T}^{k+1} \subset \mathrm{Span}\{(\varphi_{\varepsilon,T}^{k+1,i})_{1\le i\le \mathrm{N}_d^{k-1}}, (\varphi_{\varepsilon,T,F}^{k+1,j})_{F\in\mathcal{F}_T,\,1\le j\le \mathrm{N}_{d-1}^k}\},$$

since the converse inclusion follows from the definition of the oscillatory basis functions, and the cardinal of the family fits the dimension of $V_{\varepsilon,T}^{k+1}$. Let $v_\varepsilon \in V_{\varepsilon,T}^{k+1}$. Then there exist real numbers $(\theta_T^i)_{1\le i\le \mathrm{N}_d^{k-1}}$ (only if $k \ge 1$) and $(\theta_{T,F}^j)_{F\in\mathcal{F}_T,\,1\le j\le \mathrm{N}_{d-1}^k}$, satisfying the compatibility condition

$$\sum_{F\in\mathcal{F}_T}\int_F \sum_{j=1}^{\mathrm{N}_{d-1}^k} \theta_{T,F}^j \Phi_F^{k,j} = -\int_T \sum_{i=1}^{\mathrm{N}_d^{k-1}} \theta_T^i \Phi_T^{k-1,i} (= 0 \text{ if } k = 0),$$

such that

$$\begin{cases} -\mathrm{div}(\mathbb{A}_\varepsilon \nabla v_\varepsilon) = \displaystyle\sum_{i=1}^{\mathrm{N}_d^{k-1}} \theta_T^i \Phi_T^{k-1,i} (= 0 \text{ if } k = 0) & \text{in } T, \\ \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{T,F} = \displaystyle\sum_{j=1}^{\mathrm{N}_{d-1}^k} \theta_{T,F}^j \Phi_F^{k,j} & \text{on all } F \in \mathcal{F}_T. \end{cases}$$

Let us now introduce

$$\zeta := v_\varepsilon - \sum_{i=1}^{\mathrm{N}_d^{k-1}} \theta_T^i \varphi_{\varepsilon,T}^{k+1,i} - \sum_{\sigma\in\mathcal{F}_T}\sum_{j=1}^{\mathrm{N}_{d-1}^k} x_\sigma^{k,j}(v_\varepsilon)\varphi_{\varepsilon,T,\sigma}^{k+1,j},$$

where, for all $\sigma \in \mathcal{F}_T$, the real numbers $(x_\sigma^{k,j}(v_\varepsilon))_{1\le j\le \mathrm{N}_{d-1}^k}$ solve the linear system

$$\sum_{j=1}^{\mathrm{N}_{d-1}^k}\bigg(\int_\sigma \Phi_\sigma^{k,j}\,\Phi_\sigma^{k,q}\bigg) x_\sigma^{k,j}(v_\varepsilon) = \int_\sigma v_\varepsilon\,\Phi_\sigma^{k,q} \quad \text{for all } 1 \le q \le \mathrm{N}_{d-1}^k.$$

It can easily be checked that $-\mathrm{div}(\mathbb{A}_\varepsilon\nabla\zeta) = 0$ in T and that $\mathbb{A}_\varepsilon\nabla\zeta\cdot\boldsymbol{n}_{T,F} \in \mathbb{P}_{d-1}^k(F)$ and $\Pi_F^k(\zeta) = 0$ on all $F \in \mathcal{F}_T$. Using the compatibility conditions, we also infer that $\int_{\partial T} \mathbb{A}_\varepsilon\nabla\zeta\cdot\boldsymbol{n}_{\partial T} = 0$, which means that the previous system for ζ is compatible. Hence, $\zeta \equiv 0$, which concludes the proof. □

Remark 4.3 (Space $V_{\varepsilon,T}^{k+1}$). The definition of the space $V_{\varepsilon,T}^{k+1}$ is reminiscent of that considered in the nonconforming VEM in the case where $\mathbb{A}_\varepsilon = \mathbb{I}_d$; see [6] and also [14].

We define $H_{\partial T} \in \mathbb{P}_{d-1}^0(\mathcal{F}_T)$ such that, for any $F \in \mathcal{F}_T$, $H_{\partial T|F} := H_F$. We will need the following inverse inequality on the normal component of $\mathbb{A}_\varepsilon\nabla v_\varepsilon$ for a function $v_\varepsilon \in V_{\varepsilon,T}^{k+1}$; for completeness, we also establish a bound on the divergence.

Lemma 4.4 (Inverse Inequalities). *The following holds for all $v_\varepsilon \in V_{\varepsilon,T}^{k+1}$:*

$$H_T\|\mathrm{div}(\mathbb{A}_\varepsilon\nabla v_\varepsilon)\|_{L^2(T)} + \|H_{\partial T}^{\frac{1}{2}}\mathbb{A}_\varepsilon\nabla v_\varepsilon\cdot\boldsymbol{n}_{\partial T}\|_{L^2(\partial T)} \le c\,\beta^{\frac{1}{2}}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla v_\varepsilon\|_{L^2(T)^d},$$

with c independent of ε, H_T, α and β.

Proof. Note that the functions on the left-hand side are (piecewise) polynomials, but the function on the right-hand side is not a polynomial in general. Let us first bound the divergence. Let $d_\varepsilon := \mathrm{div}(\mathbb{A}_\varepsilon\nabla v_\varepsilon) \in \mathbb{P}_d^{k-1}(T)$. Let S be a simplicial sub-cell of T. Considering the standard bubble function $b_S \in H_0^1(S)$ (equal to the scaled product of the barycentric coordinates in S taking the value one at the barycenter of S), we infer using integration by parts that, for some $c > 0$ depending on mesh regularity,

$$\begin{aligned} c\,\|d_\varepsilon\|_{L^2(S)}^2 &\le \int_S d_\varepsilon b_S d_\varepsilon = \int_S \mathrm{div}(\mathbb{A}_\varepsilon\nabla v_\varepsilon) b_S d_\varepsilon \\ &= -\int_S \mathbb{A}_\varepsilon\nabla v_\varepsilon\cdot\nabla(b_S d_\varepsilon) \le \beta^{\frac{1}{2}}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla v_\varepsilon\|_{L^2(S)^d} H_S^{-1}\|d_\varepsilon\|_{L^2(S)}, \end{aligned}$$

where the last bound follows by applying an inverse inequality to the polynomial function $b_S d_\varepsilon$. Summing over all the simplicial sub-cells and invoking mesh regularity, we conclude that

$$\|\mathrm{div}(\mathbb{A}_\varepsilon \nabla v_\varepsilon)\|_{L^2(T)} \le c\,\beta^{\frac{1}{2}} H_T^{-1} \|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla v_\varepsilon\|_{L^2(T)^d}.$$

Let us now bound the normal component at the boundary. Let σ be a sub-face of a face $F \in \mathcal{F}_T$, and let $S \subseteq T$ be the simplex of the sub-mesh such that σ is a face of S. Then $r_S := [\mathrm{div}(\mathbb{A}_\varepsilon \nabla v_\varepsilon)]_{|S} \in \mathbb{P}_d^{k-1}(S) \subset \mathbb{P}_d^k(S)$ and $r_\sigma := [\mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{\partial T}]_{|\sigma} \in \mathbb{P}_{d-1}^k(\sigma)$. Note that $\boldsymbol{n}_{\partial T|\sigma} = \boldsymbol{n}_{\partial S|\sigma}$. Invoking [28, Appendix A], we infer that there is a vector-valued polynomial function $\boldsymbol{q}$ in the Raviart–Thomas–Nédélec (RTN) finite element space of order k in S so that $\mathrm{div}(\boldsymbol{q}) = r_S$ in S, $\boldsymbol{q} \cdot \boldsymbol{n}_{\partial T|\sigma} = r_\sigma$ on σ, and

$$\|\boldsymbol{q}\|_{L^2(S)^d} \le c' \min_{\substack{\boldsymbol{z} \in \boldsymbol{H}(\mathrm{div};S)\\ \mathrm{div}(\boldsymbol{z}) = r_S \text{ in } S\\ \boldsymbol{z}\cdot \boldsymbol{n}_{\partial T|\sigma} = r_\sigma \text{ on } \sigma}} \|\boldsymbol{z}\|_{L^2(S)^d},$$

with c' depending on γ (but not on k) and $\boldsymbol{H}(\mathrm{div};S) := \{\boldsymbol{z} \in L^2(S)^d : \mathrm{div}(\boldsymbol{z}) \in L^2(S)\}$. Since the function $[\mathbb{A}_\varepsilon \nabla v_\varepsilon]_{|S}$ is in $\boldsymbol{H}(\mathrm{div};S)$ and satisfies the requested conditions on the divergence in S and the normal component on σ, we conclude that $\|\boldsymbol{q}\|_{L^2(S)^d} \le c' \|\mathbb{A}_\varepsilon \nabla v_\varepsilon\|_{L^2(S)^d}$. A discrete trace inequality in the RTN finite element space shows that

$$\|\mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{\partial T}\|_{L^2(\sigma)} = \|\boldsymbol{q} \cdot \boldsymbol{n}_{\partial T}\|_{L^2(\sigma)} \le c H_\sigma^{-\frac{1}{2}} \|\boldsymbol{q}\|_{L^2(S)^d} \le c H_\sigma^{-\frac{1}{2}} \|\mathbb{A}_\varepsilon \nabla v_\varepsilon\|_{L^2(S)^d},$$

where c depends on γ and k. We conclude by invoking mesh regularity. □

4.3 Approximation Properties

We now investigate the approximation properties of the space $V_{\varepsilon,T}^{k+1}$ for all $T \in \mathcal{T}_H$. Our aim is to study how well the first-order two-scale expansion $\mathcal{L}_\varepsilon^1(u_0)$ can be approximated in the discrete space $V_{\varepsilon,T}^{k+1}$. Let us define $\pi_{\varepsilon,T}^{k+1}(u_0) \in V_{\varepsilon,T}^{k+1}$ such that $\int_T \pi_{\varepsilon,T}^{k+1}(u_0) = \int_T \mathcal{L}_\varepsilon^1(u_0)$ and

$$\begin{cases} -\mathrm{div}(\mathbb{A}_\varepsilon \nabla \pi_{\varepsilon,T}^{k+1}(u_0)) = -\mathrm{div}(\mathbb{A}_0 \nabla \Pi_T^{k+1}(u_0)) \in \mathbb{P}_d^{k-1}(T) & \text{in } T,\\ \mathbb{A}_\varepsilon \nabla \pi_{\varepsilon,T}^{k+1}(u_0) \cdot \boldsymbol{n}_{\partial T} = \mathbb{A}_0 \nabla \Pi_T^{k+1}(u_0) \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}_{d-1}^k(\mathcal{F}_T) & \text{on } \partial T. \end{cases} \tag{4.6}$$

Note that the data in (4.6) are compatible. From (4.6) we infer that, for any $w \in H^1(T)$,

$$\int_T \mathbb{A}_\varepsilon \nabla \pi_{\varepsilon,T}^{k+1}(u_0) \cdot \nabla w = \int_T \mathbb{A}_0 \nabla \Pi_T^{k+1}(u_0) \cdot \nabla w. \tag{4.7}$$

Lemma 4.5 (Approximation in $V_{\varepsilon,T}^{k+1}$). *Assume that the correctors μ_l are in $W^{1,\infty}(\mathbb{R}^d)$ for any $1 \le l \le d$, and that $u_0 \in H^{k+2}(T) \cap W^{1,\infty}(T)$. Then the following holds:*

$$\|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(\mathcal{L}_\varepsilon^1(u_0) - \pi_{\varepsilon,T}^{k+1}(u_0))\|_{L^2(T)^d} \le c\,\beta^{\frac{1}{2}} \rho^{\frac{1}{2}} \Big(H_T^{k+1} |u_0|_{H^{k+2}(T)} + \varepsilon |u_0|_{H^2(T)} + \varepsilon^{\frac{1}{2}} |\partial T|^{\frac{1}{2}} |u_0|_{W^{1,\infty}(T)} \Big), \tag{4.8}$$

with c independent of ε, H_T, u_0, α, β, and possibly depending on d, k, γ, $\max_{1 \le l \le d} \|\mu_l\|_{W^{1,\infty}(\mathbb{R}^d)}$.

Proof. Subtracting/adding $\mathbb{A}_0 \nabla u_0$ and using (4.7) with $w = \mathcal{L}_\varepsilon^1(u_0)_{|T} - \pi_{\varepsilon,T}^{k+1}(u_0)$ which is in $H^1(T)$, we infer that

$$\begin{aligned} \|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(\mathcal{L}_\varepsilon^1(u_0) - \pi_{\varepsilon,T}^{k+1}(u_0))\|_{L^2(T)^d}^2 &= \int_T (\mathbb{A}_\varepsilon \nabla \mathcal{L}_\varepsilon^1(u_0) - \mathbb{A}_0 \nabla u_0) \cdot \nabla(\mathcal{L}_\varepsilon^1(u_0) - \pi_{\varepsilon,T}^{k+1}(u_0)) \\ &\quad + \int_T \mathbb{A}_0 \nabla(u_0 - \Pi_T^{k+1}(u_0)) \cdot \nabla(\mathcal{L}_\varepsilon^1(u_0) - \pi_{\varepsilon,T}^{k+1}(u_0)). \end{aligned}$$

Using the Cauchy–Schwarz inequality and the fact that $\mathcal{L}_\varepsilon^1(u_0)_{|T} - \pi_{\varepsilon,T}^{k+1}(u_0)$ has zero mean-value on T by construction, we infer that

$$\|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(\mathcal{L}_\varepsilon^1(u_0) - \pi_{\varepsilon,T}^{k+1}(u_0))\|_{L^2(T)^d} \le \beta^{\frac{1}{2}} \rho^{\frac{1}{2}} \|\nabla(u_0 - \Pi_T^{k+1}(u_0))\|_{L^2(T)^d} + \alpha^{-\frac{1}{2}} \sup_{w \in H_\star^1(T)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\nabla w\|_{L^2(T)^d}},$$

with

$$\mathcal{F}_\varepsilon(w) = \int_T (\mathbb{A}_\varepsilon \nabla \mathcal{L}^1_\varepsilon(u_0) - \mathbb{A}_0 \nabla u_0) \cdot \nabla w \quad \text{and} \quad H^1_\star(T) = \Big\{ w \in H^1(T) : \int_T w = 0 \Big\}.$$

The first term on the right-hand side is bounded using the approximation properties (3.4) of Π_T^{k+1} with $m = 1$ and $s = k + 2$, and the second term is bounded in Lemma A.2 (take $D = T$). □

Remark 4.6 (Alternative Estimate). An alternative estimate to (4.8) can be derived under the slightly stronger regularity assumptions that there is $\kappa > 0$ so that $\mathbb{A} \in C^{0,\kappa}(\mathbb{R}^d; \mathbb{R}^{d\times d})$, and that $u_0 \in H^{\max(k+2,3)}(T)$. The proof of this estimate follows the strategy advocated in [39], where one invokes Lemma A.4 instead of Lemma A.2 at the end of the proof of Lemma 4.5 to infer that

$$\begin{aligned}\|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0))\|_{L^2(T)^d} \le c\, \beta^{\frac{1}{2}} \rho^{\frac{1}{2}} \big(H_T^{k+1} |u_0|_{H^{k+2}(T)} + (\varepsilon + (\varepsilon H_T)^{\frac{1}{2}}) |u_0|_{H^2(T)} \\ + \varepsilon H_T |u_0|_{H^3(T)} + \varepsilon^{\frac{1}{2}} H_T^{-\frac{1}{2}} |u_0|_{H^1(T)}\big),\end{aligned}$$

with c independent of ε, H_T, u_0, α, β, and possibly depending on d, k, γ, $\|\frac{\mathbb{A}}{\beta}\|_{C^{0,\kappa}(\mathbb{R}^d;\mathbb{R}^{d\times d})}$. This local estimate leads to the same global error estimate for (both versions of) the msHHO method described hereafter as (4.8); see in particular the end of the proof of Theorem 5.6.

5 The msHHO Method

In this section, we introduce and analyze the multiscale HHO (msHHO) method. We consider first in Section 5.1 a mixed-order version and then in Section 5.2 an equal-order version concerning the polynomial degree used for the cell and face unknowns. Let $\mathcal{T}_H$ be a member of an admissible mesh sequence in the sense of Definition 3.1.

5.1 The Mixed-Order Case

Let $k \ge 1$. For all $T \in \mathcal{T}_H$, we consider the following local set of discrete unknowns:

$$\underline{\mathrm{U}}_T^k := \mathbb{P}_d^{k-1}(T) \times \mathbb{P}_{d-1}^k(\mathcal{F}_T).$$

Any element $\underline{\mathrm{v}}_T \in \underline{\mathrm{U}}_T^k$ is decomposed as $\underline{\mathrm{v}}_T := (\mathrm{v}_T, \mathrm{v}_{\mathcal{F}_T})$. For any $F \in \mathcal{F}_T$, we denote $\mathrm{v}_F := \mathrm{v}_{\mathcal{F}_T|F} \in \mathbb{P}_{d-1}^k(F)$. We introduce the local reduction operator $\underline{\mathrm{I}}_T^k : H^1(T) \to \underline{\mathrm{U}}_T^k$ such that, for any $v \in H^1(T)$, $\underline{\mathrm{I}}_T^k v := (\Pi_T^{k-1}(v), \Pi_{\partial T}^k(v))$, where $\Pi_{\partial T}^k(v) \in \mathbb{P}_{d-1}^k(\mathcal{F}_T)$ is defined, for any $F \in \mathcal{F}_T$, by $\Pi_{\partial T}^k(v)_{|F} := \Pi_F^k(v)$. Reasoning as in [14, Section 2.4], it can be proved that, for all $T \in \mathcal{T}_H$, the restriction of $\underline{\mathrm{I}}_T^k$ to $V_{\varepsilon,T}^{k+1}$ is an isomorphism from $V_{\varepsilon,T}^{k+1}$ to $\underline{\mathrm{U}}_T^k$. Thus, the triple $(T, V_{\varepsilon,T}^{k+1}, \underline{\mathrm{I}}_T^k)$ defines a finite element in the sense of Ciarlet.

We define the local multiscale reconstruction operator

$$p_{\varepsilon,T}^{k+1} : \underline{\mathrm{U}}_T^k \to V_{\varepsilon,T}^{k+1}$$

such that, for any $\underline{\mathrm{v}}_T = (\mathrm{v}_T, \mathrm{v}_{\mathcal{F}_T}) \in \underline{\mathrm{U}}_T^k$, $p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) \in V_{\varepsilon,T}^{k+1}$ satisfies

$$\int_T p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) = \int_T \mathrm{v}_T$$

and solves, for all $w_\varepsilon \in V_{\varepsilon,T}^{k+1}$, the well-posed local Neumann problem

$$\int_T \mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) \cdot \nabla w_\varepsilon = -\int_T \mathrm{v}_T \operatorname{div}(\mathbb{A}_\varepsilon \nabla w_\varepsilon) + \int_{\partial T} \mathrm{v}_{\mathcal{F}_T} \mathbb{A}_\varepsilon \nabla w_\varepsilon \cdot \boldsymbol{n}_{\partial T}. \tag{5.1}$$

Note that (5.1) can be equivalently rewritten

$$\int_T \mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) \cdot \nabla w_\varepsilon = \int_T \nabla \mathrm{v}_T \cdot \mathbb{A}_\varepsilon \nabla w_\varepsilon - \int_{\partial T} (\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) \mathbb{A}_\varepsilon \nabla w_\varepsilon \cdot \boldsymbol{n}_{\partial T}. \tag{5.2}$$

Integrating by parts the left-hand side of (5.1) and exploiting the definition (4.5) of the space $V^{k+1}_{\varepsilon,T}$, one can see that, for any $\underline{\mathrm{v}}_T \in \underline{\mathrm{U}}^k_T$,

$$\Pi^{k-1}_T(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)) = \Pi^{k-1}_T(\mathrm{v}_T) = \mathrm{v}_T, \quad \Pi^k_{\partial T}(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)) = \Pi^k_{\partial T}(\mathrm{v}_{\mathcal{F}_T}) = \mathrm{v}_{\mathcal{F}_T}. \tag{5.3}$$

Owing to (4.5) and (5.1), we infer that, for all $v \in H^1(T)$,

$$\int_T \mathbb{A}_\varepsilon \nabla(v - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{I}}^k_T v)) \cdot \nabla w_\varepsilon = 0 \quad \text{for all } w_\varepsilon \in V^{k+1}_{\varepsilon,T}, \tag{5.4}$$

so that $p^{k+1}_{\varepsilon,T} \circ \underline{\mathrm{I}}^k_T : H^1(T) \to V^{k+1}_{\varepsilon,T}$ is the $\mathbb{A}_\varepsilon$-weighted elliptic projection. As a consequence, we have, for all $v \in H^1(T)$,

$$\|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(v - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{I}}^k_T v))\|_{L^2(T)^d} = \inf_{w_\varepsilon \in V^{k+1}_{\varepsilon,T}} \|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(v - w_\varepsilon)\|_{L^2(T)^d}. \tag{5.5}$$

Since the operator $p^{k+1}_{\varepsilon,T} \circ \underline{\mathrm{I}}^k_T$ preserves the mean value, its restriction to $V^{k+1}_{\varepsilon,T}$ is the identity operator.

Remark 5.1 (Comparison with the Monoscale HHO Method). In the monoscale HHO method, the reconstruction operator is simpler to construct since it maps onto $\mathbb{P}^{k+1}_d(T)$ (which is a proper subspace of $V^{k+1}_{\varepsilon,T}$ whenever $\mathbb{A}_\varepsilon$ is a constant matrix on T), whereas in the multiscale context, we explore the whole space $V^{k+1}_{\varepsilon,T}$ to build the reconstruction. One advantage of doing this is that we no longer need stabilization in the present case. Another advantage is that we recover the characterization of $p^{k+1}_{\varepsilon,T} \circ \underline{\mathrm{I}}^k_T$ as the $\mathbb{A}_\varepsilon$-weighted elliptic projector onto $V^{k+1}_{\varepsilon,T}$, that is lost in the monoscale case as soon as $\mathbb{A}_\varepsilon$ is not a constant matrix on T.

The local bilinear form $a_{\varepsilon,T} : \underline{\mathrm{U}}^k_T \times \underline{\mathrm{U}}^k_T \to \mathbb{R}$ is defined as

$$a_{\varepsilon,T}(\underline{\mathrm{u}}_T, \underline{\mathrm{v}}_T) := \int_T \mathbb{A}_\varepsilon \nabla p^{k+1}_{\varepsilon,T}(\underline{\mathrm{u}}_T) \cdot \nabla p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T).$$

We introduce the following semi-norm on $\underline{\mathrm{U}}^k_T$:

$$\|\underline{\mathrm{v}}_T\|^2_T := \|\nabla \mathrm{v}_T\|^2_{L^2(T)^d} + \|H_{\partial T}^{-\frac{1}{2}}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\|^2_{L^2(\partial T)}. \tag{5.6}$$

Lemma 5.2 (Local Stability). *The following holds:*

$$a_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T) \geq c\,\alpha \|\underline{\mathrm{v}}_T\|^2_T \quad \textit{for all } \underline{\mathrm{v}}_T \in \underline{\mathrm{U}}^k_T,$$

with constant c independent of ε, H_T, α and β.

Proof. Let $\underline{\mathrm{v}}_T \in \underline{\mathrm{U}}^k_T$. To derive an estimate on $\|\nabla \mathrm{v}_T\|_{L^2(T)^d}$, we define $v_\varepsilon \in V^{k+1}_{\varepsilon,T}$ such that

$$\begin{cases} -\operatorname{div}(\mathbb{A}_\varepsilon \nabla v_\varepsilon) = -\Delta \mathrm{v}_T \in \mathbb{P}^{k-1}_d(T) & \text{in } T, \\ \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{\partial T} = \nabla \mathrm{v}_T \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}^k_{d-1}(\mathcal{F}_T) & \text{on } \partial T, \end{cases} \tag{5.7}$$

and satisfying, e.g., $\int_T v_\varepsilon = 0$ (the way the constant is fixed is unimportant here). Note that data in (5.7) are compatible. Then the following holds:

$$\int_T \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \nabla z = \int_T \nabla \mathrm{v}_T \cdot \nabla z \quad \text{for all } z \in H^1(T).$$

Using this last relation where we take $z = p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)$, and using (5.2) where we take $w_\varepsilon = v_\varepsilon \in V^{k+1}_{\varepsilon,T}$ defined in (5.7), we infer that

$$\begin{aligned} -\int_T \mathrm{v}_T \Delta \mathrm{v}_T + \int_{\partial T} \mathrm{v}_{\mathcal{F}_T} \nabla \mathrm{v}_T \cdot \boldsymbol{n}_{\partial T} &= -\int_T \mathrm{v}_T \operatorname{div}(\mathbb{A}_\varepsilon \nabla v_\varepsilon) + \int_{\partial T} \mathrm{v}_{\mathcal{F}_T} \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{\partial T} \\ &= \int_T \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \nabla \mathrm{v}_T - \int_{\partial T} (\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \boldsymbol{n}_{\partial T} \\ &= \int_T \mathbb{A}_\varepsilon \nabla v_\varepsilon \cdot \nabla p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) = \int_T \nabla \mathrm{v}_T \cdot \nabla p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T). \end{aligned}$$

After an integration by parts, this yields

$$\|\nabla \mathrm{v}_T\|^2_{L^2(T)^d} = \int_T \nabla p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) \cdot \nabla \mathrm{v}_T + \int_{\partial T} (\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) \nabla \mathrm{v}_T \cdot \boldsymbol{n}_{\partial T}.$$

By the Cauchy–Schwarz inequality and the discrete trace inequality (3.1), we then obtain

$$\|\nabla \mathrm{v}_T\|_{L^2(T)^d} \le c\big(\alpha^{-\frac{1}{2}}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d} + \|H_{\partial T}^{-\frac{1}{2}}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\|_{L^2(\partial T)}\big). \tag{5.8}$$

To bound the second term on the right-hand side, we use (5.3) to infer that

$$\begin{aligned}[\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}]_{|\partial T} &= [\Pi_T^{k-1}(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T))]_{|\partial T} - \Pi_{\partial T}^k(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T))\\ &= \Pi_{\partial T}^k(\Pi_T^{k-1}(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)) - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)).\end{aligned}$$

Using the L^2-stability of $\Pi_{\partial T}^k$, the continuous trace inequality (3.2), the local Poincaré inequality (3.3) (since $p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) - \Pi_T^{k-1}(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T))$ has zero mean-value on T), and the H^1-stability of Π_T^{k-1}, we infer that

$$\|H_{\partial T}^{-\frac{1}{2}}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\|_{L^2(\partial T)} \le c\,\alpha^{-\frac{1}{2}}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d}. \tag{5.9}$$

This concludes the proof. □

We define the skeleton $\partial\mathcal{T}_H$ of the mesh $\mathcal{T}_H$ as $\partial\mathcal{T}_H := \bigcup_{F\in\mathcal{F}_H} F$. We introduce the broken polynomial spaces

$$\mathbb{P}_d^{k-1}(\mathcal{T}_H) := \{v \in L^2(\Omega) : v_{|T} \in \mathbb{P}_d^{k-1}(T) \text{ for all } T \in \mathcal{T}_H\}, \tag{5.10}$$
$$\mathbb{P}_{d-1}^k(\mathcal{F}_H) := \{v \in L^2(\partial\mathcal{T}_H) : v_{|F} \in \mathbb{P}_{d-1}^k(F) \text{ for all } F \in \mathcal{F}_H\}. \tag{5.11}$$

The global set of discrete unknowns is defined to be

$$\underline{\mathrm{U}}_H^k := \mathbb{P}_d^{k-1}(\mathcal{T}_H) \times \mathbb{P}_{d-1}^k(\mathcal{F}_H),$$

so that any $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k$ can be decomposed as $\underline{\mathrm{v}}_H := (\mathrm{v}_{\mathcal{T}_H}, \mathrm{v}_{\mathcal{F}_H})$. For any given discrete unknown $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k$, we denote $\underline{\mathrm{v}}_T := (\mathrm{v}_T, \mathrm{v}_{\mathcal{F}_T}) \in \underline{\mathrm{U}}_T^k$ its restriction to the mesh cell $T \in \mathcal{T}_H$. Note that unknowns attached to mesh interfaces are single-valued, in the sense that, for any $F \in \mathcal{F}_H^{\mathrm{i}}$ such that $F = \partial T_1 \cap \partial T_2 \cap Z$ for $T_1, T_2 \in \mathcal{T}_H$, $\mathrm{v}_F := \mathrm{v}_{\mathcal{F}_H|F} \in \mathbb{P}_{d-1}^k(F)$ is such that $\mathrm{v}_F = \mathrm{v}_{\mathcal{F}_{T_1}|F} = \mathrm{v}_{\mathcal{F}_{T_2}|F}$. To take into account homogeneous Dirichlet boundary conditions, we further introduce the subspace $\underline{\mathrm{U}}_{H,0}^k := \{\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k : \mathrm{v}_F \equiv 0 \text{ for all } F \in \mathcal{F}_H^{\mathrm{b}}\}$. We define the global bilinear form $a_{\varepsilon,H} : \underline{\mathrm{U}}_H^k \times \underline{\mathrm{U}}_H^k \to \mathbb{R}$ such that

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_H, \underline{\mathrm{v}}_H) := \sum_{T\in\mathcal{T}_H} a_{\varepsilon,T}(\underline{\mathrm{u}}_T, \underline{\mathrm{v}}_T) = \sum_{T\in\mathcal{T}_H} \int_T \mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{u}}_T)\cdot\nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T).$$

Then the discrete problem reads: Find $\underline{\mathrm{u}}_{\varepsilon,H} \in \underline{\mathrm{U}}_{H,0}^k$ such that

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \int_\Omega f \mathrm{v}_{\mathcal{T}_H} \quad \text{for all } \underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k. \tag{5.12}$$

Setting $\|\underline{\mathrm{v}}_H\|_H^2 := \sum_{T\in\mathcal{T}_H}\|\underline{\mathrm{v}}_T\|_T^2$ on $\underline{\mathrm{U}}_H^k$, with $\|\cdot\|_T$ introduced in (5.6), this defines a norm on $\underline{\mathrm{U}}_{H,0}^k$ since elements in $\underline{\mathrm{U}}_{H,0}^k$ are such that $\mathrm{v}_F \equiv 0$ for all $F \in \mathcal{F}_H^{\mathrm{b}}$.

Lemma 5.3 (Well-Posedness). *The following holds, for all $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k$:*

$$a_{\varepsilon,H}(\underline{\mathrm{v}}_H, \underline{\mathrm{v}}_H) = \sum_{T\in\mathcal{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d}^2 =: \|\underline{\mathrm{v}}_H\|_{\varepsilon,H}^2 \ge c\,\alpha\|\underline{\mathrm{v}}_H\|_H^2,$$

with constant c independent of ε, H, α and β. As a consequence, the discrete problem (5.12) *is well-posed.*

Proof. This is a direct consequence of Lemma 5.2. □

Remark 5.4 (Non-conforming Finite Element (ncFE) Formulation). Consider the discrete space

$$V_{\varepsilon,H,0}^{k+1} := \{v_{\varepsilon,H} \in L^2(\Omega) : v_{\varepsilon,H|T} \in V_{\varepsilon,T}^{k+1} \text{ for all } T \in \mathcal{T}_H,\ \Pi_F^k([\![v_{\varepsilon,H}]\!]_F) = 0 \text{ for all } F \in \mathcal{F}_H\},$$

where $[\![\cdot]\!]_F$ denotes the jump operator for all interfaces $F \in \mathcal{F}_H^{\mathrm{i}}$ (the sign is irrelevant) and the actual trace for all boundary faces $F \in \mathcal{F}_H^{\mathrm{b}}$. Consider the following ncFE method: Find $u_{\varepsilon,H} \in V_{\varepsilon,H,0}^{k+1}$ such that

$$\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) = \sum_{T\in\mathcal{T}_H}\int_T f\,\Pi_T^{k-1}(v_{\varepsilon,H}) \quad \text{for all } v_{\varepsilon,H} \in V_{\varepsilon,H,0}^{k+1}, \tag{5.13}$$

where

$$\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) := \sum_{T\in\mathcal{T}_H} \int_T \mathbb{A}_\varepsilon \nabla u_{\varepsilon,H} \cdot \nabla v_{\varepsilon,H}.$$

Then, using that the restriction of $\underline{\mathrm{I}}_T^k$ to $V_{\varepsilon,T}^{k+1}$ is an isomorphism from $V_{\varepsilon,T}^{k+1}$ to $\underline{\mathrm{U}}_T^k$ and that the restriction of $p_{\varepsilon,T}^{k+1} \circ \underline{\mathrm{I}}_T^k$ to $V_{\varepsilon,T}^{k+1}$ is the identity operator, it can be shown that $\underline{\mathrm{u}}_{\varepsilon,H}$ solves (5.12) if and only if $\underline{\mathrm{u}}_{\varepsilon,T} = \underline{\mathrm{I}}_T^k(u_{\varepsilon,H|T})$ for all $T \in \mathcal{T}_H$ where $u_{\varepsilon,H}$ solves (5.13). This proves that (5.12) is indeed a high-order extension of the method in [39], up to a different treatment of the right-hand side: $\Pi_T^{k-1}(v_{\varepsilon,H})$ is used instead of $v_{\varepsilon,H}$.

Let u_ε be the oscillatory solution to (1.1) and let $\underline{\mathrm{u}}_{\varepsilon,H}$ be the discrete msHHO solution to (5.12). Let us define the discrete error such that

$$\underline{\mathrm{e}}_{\varepsilon,H} \in \underline{\mathrm{U}}_{H,0}^k, \quad \underline{\mathrm{e}}_{\varepsilon,T} := \underline{\mathrm{I}}_T^k u_\varepsilon - \underline{\mathrm{u}}_{\varepsilon,T} \quad \text{for all } T \in \mathcal{T}_H. \tag{5.14}$$

Note that $\underline{\mathrm{e}}_{\varepsilon,H}$ is well-defined as a member of $\underline{\mathrm{U}}_{H,0}^k$ since the oscillatory solution u_ε is in $H_0^1(\Omega)$ and functions in $H_0^1(\Omega)$ are single-valued at interfaces and vanish at the boundary.

Lemma 5.5 (Discrete Energy-Error Estimate). *Let the discrete error $\underline{\mathrm{e}}_{\varepsilon,H}$ be defined by* (5.14). *Assume that $u_0 \in H^{k+2}(\Omega)$. Then the following holds:*

$$\|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H} \le c\,\rho^{\frac{1}{2}}\Big(\beta \sum_{T\in\mathcal{T}_H} H_T^{2(k+1)} |u_0|^2_{H^{k+2}(T)} + \sum_{T\in\mathcal{T}_H} \|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(u_\varepsilon - \pi_{\varepsilon,T}^{k+1}(u_0))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}}, \tag{5.15}$$

with constant c independent of ε, H, u_0, α and β.

Proof. Lemma 5.3 implies that

$$\|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H} = \sup_{\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k} \frac{a_{\varepsilon,H}(\underline{\mathrm{e}}_{\varepsilon,H}, \underline{\mathrm{v}}_H)}{\|\underline{\mathrm{v}}_H\|_{\varepsilon,H}}.$$

Let $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k$. Performing an integration by parts, and using the facts that the flux $\mathbb{A}_0 \nabla u_0 \cdot \boldsymbol{n}_F$ is continuous across any interface $F \in \mathcal{F}_H^{\mathrm{i}}$ since $u_0 \in H^2(\Omega)$, and that $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k$, we infer that

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \int_\Omega f \mathrm{v}_{\mathcal{T}_H} = \sum_{T\in\mathcal{T}_H} \int_T \mathbb{A}_0 \nabla u_0 \cdot \nabla \mathrm{v}_T - \sum_{T\in\mathcal{T}_H} \int_{\partial T} (\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) \mathbb{A}_0 \nabla u_0 \cdot \boldsymbol{n}_{\partial T}.$$

Using (5.2) with $w_\varepsilon = p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon)$, we then infer that

$$a_{\varepsilon,H}(\underline{\mathrm{e}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \sum_{T\in\mathcal{T}_H} \int_T (\mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \mathbb{A}_0 \nabla u_0) \cdot \nabla \mathrm{v}_T - \sum_{T\in\mathcal{T}_H} \int_{\partial T} (\mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \mathbb{A}_0 \nabla u_0) \cdot \boldsymbol{n}_{\partial T}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}).$$

Adding/subtracting $\Pi_T^{k+1}(u_0)$ on the right-hand side yields $a_{\varepsilon,H}(\underline{\mathrm{e}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \mathfrak{T}_1 + \mathfrak{T}_2$ with

$$\mathfrak{T}_1 = \sum_{T\in\mathcal{T}_H} \int_T \mathbb{A}_0 \nabla(\Pi_T^{k+1}(u_0) - u_0) \cdot \nabla \mathrm{v}_T - \sum_{T\in\mathcal{T}_H} \int_{\partial T} \mathbb{A}_0 \nabla(\Pi_T^{k+1}(u_0) - u_0) \cdot \boldsymbol{n}_{\partial T}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}),$$

$$\mathfrak{T}_2 = \sum_{T\in\mathcal{T}_H} \int_T \big(\mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \mathbb{A}_0 \nabla \Pi_T^{k+1}(u_0)\big) \cdot \nabla \mathrm{v}_T - \sum_{T\in\mathcal{T}_H} \int_{\partial T} \big(\mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \mathbb{A}_0 \nabla \Pi_T^{k+1}(u_0)\big) \cdot \boldsymbol{n}_{\partial T}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}).$$

The term $\mathfrak{T}_1$ is estimated using Cauchy–Schwarz inequality and the approximation properties (3.4) of the projector Π_T^{k+1} for $m = 1$ and $s = k + 2$, yielding

$$|\mathfrak{T}_1| \le c\,\beta\Big(\sum_{T\in\mathcal{T}_H} H_T^{2(k+1)} |u_0|^2_{H^{k+2}(T)}\Big)^{\frac{1}{2}} \|\underline{\mathrm{v}}_H\|_H.$$

Considering now $\mathfrak{T}_2$, we use the definition (4.6) of $\pi_{\varepsilon,T}^{k+1}(u_0)$ and relation (4.7) to infer that

$$\mathfrak{T}_2 = \sum_{T\in\mathcal{T}_H} \int_T \mathbb{A}_\varepsilon \nabla(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \pi_{\varepsilon,T}^{k+1}(u_0)) \cdot \nabla \mathrm{v}_T - \sum_{T\in\mathcal{T}_H} \int_{\partial T} \mathbb{A}_\varepsilon \nabla(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \pi_{\varepsilon,T}^{k+1}(u_0)) \cdot \boldsymbol{n}_{\partial T}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}).$$

The first term on the right-hand side can be bounded using the Cauchy–Schwarz inequality, whereas the second term is estimated by means of the inverse inequality from Lemma 4.4 since $(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{I}}^k_T u_\varepsilon) - \pi^{k+1}_{\varepsilon,T}(u_0)) \in V^{k+1}_{\varepsilon,T}$. This yields

$$
\begin{aligned}
|\mathfrak{T}_2| &\leq c\,\beta^{\frac{1}{2}}\Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{I}}^k_T u_\varepsilon) - \pi^{k+1}_{\varepsilon,T}(u_0))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}}\|\underline{\mathrm{v}}_H\|_H \\
&\leq c\,\beta^{\frac{1}{2}}\Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \pi^{k+1}_{\varepsilon,T}(u_0))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}}\|\underline{\mathrm{v}}_H\|_H,
\end{aligned}
$$

where the last bound follows from (5.5) since $\pi^{k+1}_{\varepsilon,T}(u_0) \in V^{k+1}_{\varepsilon,T}$. Since $\|\underline{\mathrm{v}}_H\|^2_{\varepsilon,H} \geq c\,\alpha\|\underline{\mathrm{v}}_H\|^2_H$ owing to Lemma 5.3, we obtain the expected bound. □

Theorem 5.6 (Energy-Error Estimate). *Assume that the correctors μ_l are in $W^{1,\infty}(\mathbb{R}^d)$ for any integer $1 \leq l \leq d$, and that $u_0 \in H^{k+2}(\Omega)$ (recall that $k \geq 1$). Then the following holds:*

$$
\begin{aligned}
\Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{u}}_{\varepsilon,T}))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} &\leq c\,\beta^{\frac{1}{2}}\rho\Big(\sum_{T\in\mathfrak{T}_H}H_T^{2(k+1)}|u_0|^2_{H^{k+2}(T)} + \varepsilon|\partial\Omega||u_0|^2_{W^{1,\infty}(\Omega)} \\
&\quad + \sum_{T\in\mathfrak{T}_H}[\varepsilon^2|u_0|^2_{H^2(T)} + \varepsilon|\partial T||u_0|^2_{W^{1,\infty}(T)}]\Big)^{\frac{1}{2}},
\end{aligned}
\tag{5.16}
$$

with c independent of ε, H, u_0, α and β. In particular, if the mesh $\mathfrak{T}_H$ is quasi-uniform, and tracking for simplicity only the dependency on ε and H with $\varepsilon \leq H \leq \ell_\Omega$ (ℓ_Ω denotes the diameter of Ω), we obtain an energy-error upper bound of the form $(\varepsilon^{\frac{1}{2}} + H^{k+1} + (\frac{\varepsilon}{H})^{\frac{1}{2}})$.

Proof. Using the shorthand notation $e_{\varepsilon,T} := u_\varepsilon|_T - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{u}}_{\varepsilon,T})$ for all $T \in \mathfrak{T}_H$, the triangle inequality implies that

$$
\Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla e_{\varepsilon,T}\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} \leq \Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{I}}^k_T u_\varepsilon))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} + \|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H},
$$

and owing to (5.5), we infer that

$$
\Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla e_{\varepsilon,T}\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} \leq \Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \pi^{k+1}_{\varepsilon,T}(u_0))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} + \|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H}.
$$

Lemma 5.5 then implies that

$$
\Big(\sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla e_{\varepsilon,T}\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} \leq c\,\rho^{\frac{1}{2}}\Big(\beta\sum_{T\in\mathfrak{T}_H}H_T^{2(k+1)}|u_0|^2_{H^{k+2}(T)} + \sum_{T\in\mathfrak{T}_H}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \pi^{k+1}_{\varepsilon,T}(u_0))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}}.
$$

To conclude the proof of (5.16), we add/subtract $\mathcal{L}^1_\varepsilon(u_0)$ in the last term on the right-hand side, and invoke the triangle inequality together with Lemma A.5 to bound $(u_\varepsilon - \mathcal{L}^1_\varepsilon(u_0))$ globally on Ω and Lemma 4.5 to bound $(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0))$ locally on all $T \in \mathfrak{T}_H$. Finally, to derive the upper bound for quasi-uniform meshes, we observe that the last term in (5.16) can be estimated as

$$
\sum_{T\in\mathfrak{T}_H}\varepsilon|\partial T||u_0|^2_{W^{1,\infty}(T)} \leq c\,\varepsilon H^{-1}|u_0|^2_{W^{1,\infty}(\Omega)}\sum_{T\in\mathfrak{T}_H}|\partial T|H_T \leq c'\varepsilon H^{-1}|u_0|^2_{W^{1,\infty}(\Omega)}
$$

with c' proportional to $|\Omega|$. □

Remark 5.7 (Dependency on ρ). Estimate (5.16) has a linear dependency with respect to the (global) heterogeneity/anisotropy ratio ρ (a close inspection of the proof shows that the term $\varepsilon^{\frac{1}{2}}|\partial\Omega|^{\frac{1}{2}}|u_0|_{W^{1,\infty}(\Omega)}$ only scales with $\rho^{\frac{1}{2}}$). This linear scaling is also obtained with the monoscale HHO method when the diffusivity is non-constant in each mesh cell; cf. [20, Theorem 3.1].

Remark 5.8 (Discretization of the Right-Hand Side). Note that we could also integrate the right-hand side in (5.12) using $p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)$ instead of v_T on each $T \in \mathfrak{T}_H$, up to the addition on the right-hand sides of the bounds (5.15) and (5.16) of the optimally convergent term $c\,\alpha^{-\frac{1}{2}}(\sum_{T\in\mathcal{T}_H} H_T^{2(k+1)}|f|^2_{H^k(T)})^{\frac{1}{2}}$. Indeed, owing to (5.3), we have

$$\sum_{T\in\mathcal{T}_H}\int_T f\,(\mathrm{v}_T - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)) = \sum_{T\in\mathcal{T}_H}\int_T (f-\Pi_T^{k-1}(f))\,(\mathrm{v}_T - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)),$$

which can be estimated by applying the Cauchy–Schwarz inequality on each T, and

(i) the approximation properties (3.4) of Π_T^{k-1} with $m = 0$ and $s = k$ for the first factor,

(ii) the Poincaré inequality (3.3) (recall that $(\mathrm{v}_T - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T))$ has zero-mean on T) and the triangle inequality combined with Lemma 5.2 for the second factor.

This alternative approach, that is pursued in [39, 40], necessitates an integration against oscillatory test functions. It is hence computationally more expensive (recall that f is assumed to be non-oscillatory), and may become limiting in a multi-query context.

5.2 The Equal-Order Case

Let $k \geq 0$. For all $T \in \mathcal{T}_H$, we consider now the following local set of discrete unknowns:

$$\underline{\mathrm{U}}_T^k := \mathbb{P}_d^k(T) \times \mathbb{P}_{d-1}^k(\mathcal{F}_T).$$

Any $\underline{\mathrm{v}}_T \in \underline{\mathrm{U}}_T^k$ is again decomposed as $\underline{\mathrm{v}}_T := (\mathrm{v}_T, \mathrm{v}_{\mathcal{F}_T})$, and for any $F \in \mathcal{F}_T$, we denote $\mathrm{v}_F := \mathrm{v}_{\mathcal{F}_T|F} \in \mathbb{P}_{d-1}^k(F)$. We redefine the local reduction operator $\underline{\mathrm{I}}_T^k : H^1(T) \to \underline{\mathrm{U}}_T^k$ so that, for any $v \in H^1(T)$,

$$\underline{\mathrm{I}}_T^k v := (\Pi_T^k(v), \Pi_{\partial T}^k(v)).$$

Reasoning as in [14, Section 2.4], it can be proved that, for all $T \in \mathcal{T}_H$, the restriction of $\underline{\mathrm{I}}_T^k$ to $\tilde{V}^{k+1}_{\varepsilon,T}$ is an isomorphism from $\tilde{V}^{k+1}_{\varepsilon,T}$ to $\underline{\mathrm{U}}_T^k$, where

$$\tilde{V}^{k+1}_{\varepsilon,T} := \{v_\varepsilon \in H^1(T) : \operatorname{div}(\mathbb{A}_\varepsilon\nabla v_\varepsilon) \in \mathbb{P}_d^k(T),\ \mathbb{A}_\varepsilon\nabla v_\varepsilon\cdot\boldsymbol{n}_{\partial T} \in \mathbb{P}_{d-1}^k(\mathcal{F}_T)\}.$$

Thus, the triple $(T, \tilde{V}^{k+1}_{\varepsilon,T}, \underline{\mathrm{I}}_T^k)$ defines a finite element in the sense of Ciarlet.

The local multiscale reconstruction operator $p^{k+1}_{\varepsilon,T} : \underline{\mathrm{U}}_T^k \to V^{k+1}_{\varepsilon,T}$ is still defined as in (5.1), so that the key relations (5.4) and (5.5) still hold. In particular, $p^{k+1}_{\varepsilon,T}\circ\underline{\mathrm{I}}_T^k : H^1(T) \to V^{k+1}_{\varepsilon,T}$ is the $\mathbb{A}_\varepsilon$-weighted elliptic projection. However, the restriction of $p^{k+1}_{\varepsilon,T}\circ\underline{\mathrm{I}}_T^k$ to the larger space $\tilde{V}^{k+1}_{\varepsilon,T}$ is *not* the identity operator since $p^{k+1}_{\varepsilon,T}$ maps onto the smaller space $V^{k+1}_{\varepsilon,T}$. Concerning (5.3), we still have

$$\Pi_{\partial T}^k(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)) = \mathrm{v}_{\mathcal{F}_T},$$

but now $\Pi_T^{k-1}(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)) = \Pi_T^{k-1}(\mathrm{v}_T)$ is in general different from v_T.

This leads us to introduce the symmetric, positive semi-definite stabilization

$$j_{\varepsilon,T}(\underline{\mathrm{u}}_T, \underline{\mathrm{v}}_T) := \alpha\int_{\partial T} H_{\partial T}^{-1}(\mathrm{u}_T - \Pi_T^k(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{u}}_T)))(\mathrm{v}_T - \Pi_T^k(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T))).$$

The local bilinear form $a_{\varepsilon,T} : \underline{\mathrm{U}}_T^k \times \underline{\mathrm{U}}_T^k \to \mathbb{R}$ is then defined as

$$a_{\varepsilon,T}(\underline{\mathrm{u}}_T, \underline{\mathrm{v}}_T) := \int_T \mathbb{A}_\varepsilon\nabla p^{k+1}_{\varepsilon,T}(\underline{\mathrm{u}}_T)\cdot\nabla p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) + j_{\varepsilon,T}(\underline{\mathrm{u}}_T, \underline{\mathrm{v}}_T).$$

Remark 5.9 (Variant). Alternatively, one can discard the stabilization at the price of computing additional cell-based oscillatory basis functions, using the basis functions $(\Phi_T^{k,i})_{1\leq i\leq \mathrm{N}_d^k}$ instead of $(\Phi_T^{k-1,i})_{1\leq i\leq \mathrm{N}_d^{k-1}}$ as proposed in Section 4.1.1. This is the approach pursued in [40] for $k = 0$ where one cell-based oscillatory basis function is added (in the slightly different context of perforated domains).

Recall the local stability semi-norm $\|\cdot\|_T$ defined by (5.6).

Lemma 5.10 (Local Stability and Approximation). *The following holds:*

$$a_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T) \geq c\,\alpha \|\underline{\mathrm{v}}_T\|_T^2 \quad \text{for all } \underline{\mathrm{v}}_T \in \underline{\mathrm{U}}_T^k.$$

Moreover, for all $v \in H^1(T)$,

$$j_{\varepsilon,T}(\underline{\mathrm{I}}_T^k v, \underline{\mathrm{I}}_T^k v)^{\frac{1}{2}} \leq c\,\|\mathbb{A}_\varepsilon^{\frac{1}{2}} \nabla(v - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k v))\|_{L^2(T)^d}, \tag{5.17}$$

with (distinct) constants c independent of ε, H_T, α and β.

Proof. To prove stability, we adapt the proof of Lemma 5.2. Let $\underline{\mathrm{v}}_T \in \underline{\mathrm{U}}_T^k$. The bound (5.8) on $\|\nabla \mathrm{v}_T\|_{L^2(T)^d}$ still holds, so that we only need to bound $\|H_{\partial T}^{-1/2}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\|_{L^2(\partial T)}$. Since $\Pi_{\partial T}^k(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)) = \mathrm{v}_{\mathcal{F}_T}$, we infer that

$$(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) = \Pi_{\partial T}^k(\mathrm{v}_T - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)),$$

so that invoking the L^2-stability of $\Pi_{\partial T}^k$ and the triangle inequality while adding/subtracting $\Pi_T^k(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T))$, we obtain

$$\|H_{\partial T}^{-\frac{1}{2}}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\|_{L^2(\partial T)} \leq \|H_{\partial T}^{-\frac{1}{2}}(\mathrm{v}_T - \Pi_T^k(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)))\|_{L^2(\partial T)} + \|H_{\partial T}^{-\frac{1}{2}}(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) - \Pi_T^k(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)))\|_{L^2(\partial T)}.$$

The first term on the right-hand side is bounded by $\alpha^{-\frac{1}{2}} j_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T)^{\frac{1}{2}}$, and the second one has been bounded (with the use of Π_T^{k-1} instead of Π_T^k) in the proof of Lemma 5.2 (see (5.9)) by $c\,\alpha^{-1/2}\|\mathbb{A}_\varepsilon^{1/2}\nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d}$. To prove (5.17), we start from

$$j_{\varepsilon,T}(\underline{\mathrm{I}}_T^k v, \underline{\mathrm{I}}_T^k v) = \alpha \|H_{\partial T}^{-\frac{1}{2}} \Pi_T^k(v - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k v))\|_{L^2(\partial T)}^2.$$

The result then follows from the application of the discrete trace inequality (3.1), the L^2-stability property of Π_T^k, and the local Poincaré inequality (3.3) (since $\int_T p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k v) = \int_T v$). □

We define the broken polynomial space

$$\mathbb{P}_d^k(\mathcal{T}_H) := \{v \in L^2(\Omega) : v_{|T} \in \mathbb{P}_d^k(T) \text{ for all } T \in \mathcal{T}_H\},$$

and the global set of discrete unknowns is defined to be

$$\underline{\mathrm{U}}_H^k := \mathbb{P}_d^k(\mathcal{T}_H) \times \mathbb{P}_{d-1}^k(\mathcal{F}_H),$$

where $\mathbb{P}_{d-1}^k(\mathcal{F}_H)$ is still defined by (5.11). To take into account homogeneous Dirichlet boundary conditions, we consider again the subspace $\underline{\mathrm{U}}_{H,0}^k := \{\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k : \mathrm{v}_F \equiv 0 \text{ for all } F \in \mathcal{F}_H^{\mathrm{b}}\}$. We define the global bilinear form $a_{\varepsilon,H} : \underline{\mathrm{U}}_H^k \times \underline{\mathrm{U}}_H^k \to \mathbb{R}$ such that

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_H, \underline{\mathrm{v}}_H) := \sum_{T\in\mathcal{T}_H} a_{\varepsilon,T}(\underline{\mathrm{u}}_T, \underline{\mathrm{v}}_T) = \sum_{T\in\mathcal{T}_H} \bigg(\int_T \mathbb{A}_\varepsilon \nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{u}}_T)\cdot\nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) + j_{\varepsilon,T}(\underline{\mathrm{u}}_T, \underline{\mathrm{v}}_T)\bigg).$$

Then the discrete problem reads: Find $\underline{\mathrm{u}}_{\varepsilon,H} \in \underline{\mathrm{U}}_{H,0}^k$ such that

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \int_\Omega f \mathrm{v}_{\mathcal{T}_H} \quad \text{for all } \underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k. \tag{5.18}$$

Recalling the norm $\|\underline{\mathrm{v}}_H\|_H^2 := \sum_{T\in\mathcal{T}_H}\|\underline{\mathrm{v}}_T\|_T^2$ on $\underline{\mathrm{U}}_{H,0}^k$, we readily infer from Lemma 5.10 the following well-posedness result.

Lemma 5.11 (Well-Posedness). *The following holds, for all $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k$:*

$$a_{\varepsilon,H}(\underline{\mathrm{v}}_H, \underline{\mathrm{v}}_H) = \sum_{T\in\mathcal{T}_H} (\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d}^2 + j_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T)) =: \|\underline{\mathrm{v}}_H\|_{\varepsilon,H}^2 \geq c\,\alpha\|\underline{\mathrm{v}}_H\|_H^2,$$

with constant c independent of ε, H, α and β. As a consequence, the discrete problem (5.18) *is well-posed.*

Remark 5.12 (ncFE Interpretation). As in Remark 5.4, it is possible to give a ncFE interpretation of the scheme (5.18). Let

$$\tilde{V}_{\varepsilon,H,0}^{k+1} := \{v_{\varepsilon,H} \in L^2(\Omega) : v_{\varepsilon,H|T} \in \tilde{V}_{\varepsilon,T}^{k+1} \text{ for all } T \in \mathcal{T}_H,\ \Pi_F^k([\![v_{\varepsilon,H}]\!]_F) = 0 \text{ for all } F \in \mathcal{F}_H\},$$

and consider the following ncFE method: Find $u_{\varepsilon,H} \in \tilde{V}^{k+1}_{\varepsilon,H,0}$ such that

$$\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) = \sum_{T\in\mathcal{T}_H} \int_T f\,\Pi^k_T(v_{\varepsilon,H}) \quad \text{for all } v_{\varepsilon,H} \in \tilde{V}^{k+1}_{\varepsilon,H,0}, \tag{5.19}$$

where

$$\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) := \sum_{T\in\mathcal{T}_H} a_{\varepsilon,T}(\underline{\mathrm{I}}^k_T(u_{\varepsilon,H|T}), \underline{\mathrm{I}}^k_T(v_{\varepsilon,H|T})).$$

Then it can be shown that $\underline{\mathrm{u}}_{\varepsilon,H}$ solves (5.18) if and only if $\underline{\mathrm{u}}_{\varepsilon,T} = \underline{\mathrm{I}}^k_T(u_{\varepsilon,H|T})$ for all $T \in \mathcal{T}_H$, where $u_{\varepsilon,H}$ solves (5.19). The main difference with respect to the mixed-order case is that it is no longer possible to simplify the expression of the bilinear form $\tilde{a}_{\varepsilon,H}$ since the restriction of $p^{k+1}_{\varepsilon,T} \circ \underline{\mathrm{I}}^k_T$ to $\tilde{V}^{k+1}_{\varepsilon,T}$ is not the identity operator. As in the monoscale HHO method, the operator $p^{k+1}_{\varepsilon,T}$, which maps onto the smaller space $V^{k+1}_{\varepsilon,T}$, allows one to restrict the number of computed basis functions while maintaining optimal (and here also ε-robust) approximation properties. The functions (from the discrete space $\tilde{V}^{k+1}_{\varepsilon,T}$) that are eliminated (not computed) are handled by the stabilization term.

Lemma 5.13 (Discrete Energy-Error Estimate). *Let the discrete error $\underline{\mathrm{e}}_{\varepsilon,H}$ be defined by* (5.14). *Assume that $u_0 \in H^{k+2}(\Omega)$. Then the following holds:*

$$\|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H} \le c\,\rho^{\frac{1}{2}}\Big(\beta \sum_{T\in\mathcal{T}_H} H_T^{2(k+1)}|u_0|^2_{H^{k+2}(T)} + \sum_{T\in\mathcal{T}_H} \|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \pi^{k+1}_{\varepsilon,T}(u_0))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}},$$

with constant c independent of ε, H, u_0, α and β.

Proof. The only difference with the proof of Lemma 5.5 is that we now have $a_{\varepsilon,H}(\underline{\mathrm{e}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \mathfrak{T}_1 + \mathfrak{T}_2 + \mathfrak{T}_3$, where $\mathfrak{T}_1, \mathfrak{T}_2$ are defined and bounded in that proof and where

$$\mathfrak{T}_3 := \sum_{T\in\mathcal{T}_H} j_{\varepsilon,T}(\underline{\mathrm{I}}^k_T u_\varepsilon, \underline{\mathrm{v}}_T).$$

Since $j_{\varepsilon,T}$ is symmetric, positive semi-definite, we infer that

$$\begin{aligned}|\mathfrak{T}_3| &\le \Big(\sum_{T\in\mathcal{T}_H} j_{\varepsilon,T}(\underline{\mathrm{I}}^k_T u_\varepsilon, \underline{\mathrm{I}}^k_T u_\varepsilon)\Big)^{\frac{1}{2}}\Big(\sum_{T\in\mathcal{T}_H} j_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T)\Big)^{\frac{1}{2}} \\ &\le c\Big(\sum_{T\in\mathcal{T}_H} \|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{I}}^k_T u_\varepsilon))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} \|\underline{\mathrm{v}}_H\|_{\varepsilon,H},\end{aligned}$$

where we have used (5.17). We can now conclude as before. □

Theorem 5.14 (Energy-Error Estimate). *Assume that the correctors μ_l are in $W^{1,\infty}(\mathbb{R}^d)$ for any $1 \le l \le d$, and that $u_0 \in H^{k+2}(\Omega) \cap W^{1,\infty}(\Omega)$. Then the following holds:*

$$\begin{aligned}\Big(\sum_{T\in\mathcal{T}_H} \|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{u}}_{\varepsilon,T}))\|^2_{L^2(T)^d}\Big)^{\frac{1}{2}} \le c\,\beta^{\frac{1}{2}}\rho\Big(&\sum_{T\in\mathcal{T}_H} H_T^{2(k+1)}|u_0|^2_{H^{k+2}(T)} + \varepsilon|\partial\Omega||u_0|^2_{W^{1,\infty}(\Omega)} \\ &+ \sum_{T\in\mathcal{T}_H}[\varepsilon^2|u_0|^2_{H^2(T)} + \varepsilon|\partial T||u_0|^2_{W^{1,\infty}(T)}]\Big)^{\frac{1}{2}},\end{aligned} \tag{5.20}$$

with c independent of ε, H, u_0, α and β. In particular, if the mesh $\mathcal{T}_H$ is quasi-uniform, and tracking for simplicity only the dependency on ε and H with $\varepsilon \le H \le \ell_\Omega$, we obtain an energy-error upper bound of the form $(\varepsilon^{\frac{1}{2}} + H^{k+1} + (\frac{\varepsilon}{H})^{\frac{1}{2}})$.

Proof. Identical to that of Theorem 5.6. □

Remark 5.15 (Dependency on ρ). As in the mixed-order case (cf. Remark 5.7), estimate (5.20) has a linear dependency with respect to the (global) heterogeneity/anisotropy ratio ρ.

Remark 5.16 (Discretization of the Right-Hand Side). The same observation as in Remark 5.8 concerning the discretization of the right-hand side in (5.18) is still valid for the equal-order case.

6 Numerical Results

In this section, we discuss the organization of the computations and we present some numerical results illustrating the above analysis for both the mixed-order and equal-order msHHO methods. Our numerical results have been obtained using the `Disk++` library, which is available as open-source under MPL license at the address https://github.com/datafl4sh/diskpp. The numerical core of the library is described in [12]. For the numerical tests presented below, we have used the direct solver PARDISO of the Intel MKL library. The simulations were run on an Intel i7-3615QM (2.3 GHz) with 16 Gb of RAM.

6.1 Offline/Online Solution Strategy

Let us consider the equal-order version ($k \geq 0$) of the msHHO method introduced in Section 5.2. Similar considerations carry over to the mixed-order case ($k \geq 1$) of Section 5.1. To solve problem (5.18), we adopt an offline/online strategy.

- In the offline step, all the computations are local, and independent of the right-hand side f. We first compute the cell-based and face-based basis functions, i.e., for all $T \in \mathcal{T}_H$, we compute the N_d^{k-1} functions $\varphi_{\varepsilon,T}^{k+1,i}$ solution to (4.2), and the $\mathrm{card}(\mathcal{F}_T) \times \mathrm{N}_{d-1}^k$ functions $\varphi_{\varepsilon,T,F}^{k+1,j}$ solution to (4.4) (cf. Remark 4.1). This first substep is fully parallelizable. In a second time, we compute the multiscale reconstruction operators $p_{\varepsilon,T}^{k+1}$, by solving (5.1) for all $T \in \mathcal{T}_H$. Each computation requires to invert a symmetric positive-definite matrix of size $(\mathrm{N}_d^{k-1} + \mathrm{card}(\mathcal{F}_T) \times \mathrm{N}_{d-1}^k)$, which can be performed effectively via Cholesky factorization. This second substep is as well fully parallelizable. Finally, we perform static condensation locally in each cell of $\mathcal{T}_H$, to eliminate the cell unknowns. Details can be found in [20, Section 3.3.1]. Basically, in each cell, this substep consists in inverting a symmetric positive-definite matrix of size N_d^k. This last substep is also fully parallelizable.
- In the online step, we compute the L^2-orthogonal projection of the right-hand side f onto $\mathbb{P}_d^k(\mathcal{T}_H)$, and we then solve a symmetric positive-definite global problem, posed in terms of the face unknowns only. The size of this problem is $\mathrm{card}(\mathcal{F}_H^{\mathrm{i}}) \times \mathrm{N}_{d-1}^k$. If one wants to compute an approximation of the solution to (1.1) for another f (or for other boundary conditions), only the online step must be rerun.

6.2 Periodic Test-Case

We consider the periodic test-case studied in [39] (and also in [43]). We let $d = 2$, and let Ω be the unit square. We consider problem (1.1), with right-hand side $f(x, y) = \sin(x)\sin(y)$, and oscillatory coefficient

$$\mathbb{A}_\varepsilon(x, y) = a(\tfrac{x}{\varepsilon}, \tfrac{y}{\varepsilon})\mathbb{I}_2, \quad a(x_1, x_2) = 1 + 100\cos^2(\pi x_1)\sin^2(\pi x_2). \tag{6.1}$$

For the coefficient (6.1), the homogenized tensor is given by $\mathbb{A}_0 \approx 6.72071\,\mathbb{I}_2$. We fix $\varepsilon = \frac{\pi}{150} \approx 0.021$.

We consider a sequence of hierarchical triangular meshes of size $H_l = 0.43 \times 2^{-l}$ with $l \in \{0, \dots, 9\}$, so that $H_5 < \varepsilon < H_4$. A reference solution is computed by solving (1.1) with the (equal-order) monoscale HHO method on the mesh of level $l_{\mathrm{ref}} = 9$ with polynomial degree $k_{\mathrm{ref}} = 2$. In Figure 1, we present the (absolute) energy-norm errors obtained with the msHHO method on the meshes $\mathcal{T}_{H_l}$ with $l \in \{0, \dots, 6\}$. We consider both the mixed-order msHHO method with polynomial degrees $k \in \{1, 2\}$ and the equal-order msHHO method with polynomial degrees $k \in \{0, 1, 2\}$. In all cases, the cell- and face-based oscillatory basis functions are precomputed using the (equal-order) monoscale HHO method on the mesh of level $l_{\mathrm{osc}} = 8$ with polynomial degree $k_{\mathrm{osc}} = 1$. We have verified that the oscillatory basis functions are sufficiently well resolved by comparing our results to those obtained with $k_{\mathrm{osc}} = 2$ and obtaining only very marginal differences. The first observation we draw from Figure 1 is that the mixed-order and equal-order msHHO methods employing the same polynomial degree for the face unknowns deliver very similar results; indeed, the error curves are barely distinguishable both for $k = 1$ and $k = 2$. Moreover, we can observe all the main features expected from the error analysis: a pre-asymptotic regime where the term H^{k+1} essentially dominates (meshes of levels $l \in \{0, 1\}$),

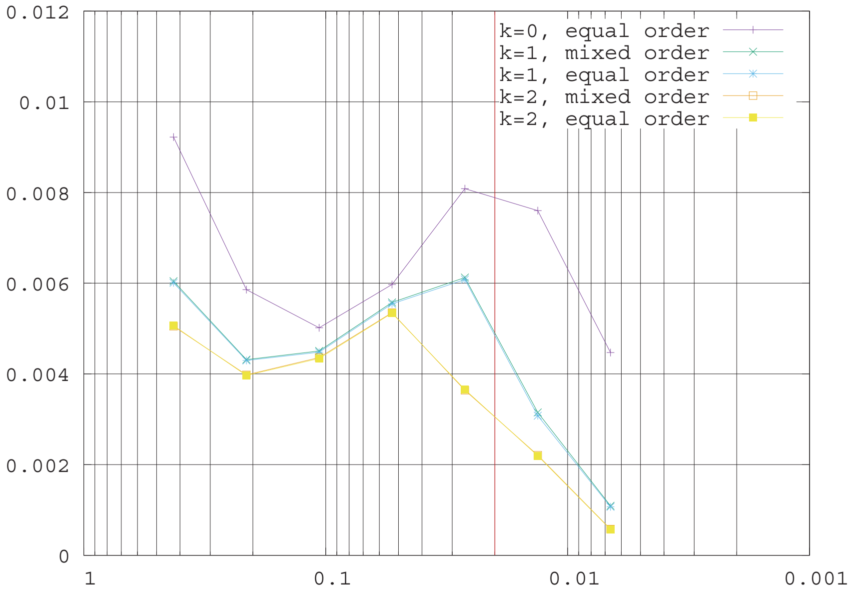

Figure 1: Periodic test-case: convergence results in energy-norm for mesh levels $l \in \{0, \dots, 6\}$; mixed-order msHHO method with polynomial degrees $k \in \{1, 2\}$ and equal-order msHHO method with polynomial degrees $k \in \{0, 1, 2\}$. The red vertical line indicates the value of ε.

the resonance regime (meshes of levels $l \in \{2, 3, 4\}$ essentially), and the asymptotic regime where the mesh actually resolves the fine scale of the model coefficients (meshes of levels $l \in \{5, 6\}$). We can also see the advantages of using a higher polynomial order for the face unknowns: the error is overall smaller, the minimal error in the resonance regime is reached at a larger value of H and takes a smaller value (incidentally, the maximal error in the resonance regime takes a smaller value as well), and the asymptotic regime starts for larger values of H.

6.3 Locally Periodic Test-Case

Keeping the same two-dimensional domain Ω as in the periodic test-case of Section 6.2, we consider now a locally periodic test-case where we solve problem (1.1) with unchanged right-hand side $f(x, y) = \sin(x)\sin(y)$, but with oscillatory coefficient

$$\mathbb{A}_\varepsilon(x, y) = \big(a(\tfrac{x}{\varepsilon}, \tfrac{y}{\varepsilon}) + e^{\frac{x^2+y^2}{2}}\big)\mathbb{I}_2, \quad \text{with } a \text{ given in (6.1)},$$

and with unchanged value of ε. We perform the same numerical experiments as in Section 6.2 using the same mesh level and polynomial order parameters for computing the reference solution and the oscillatory basis functions (we verified similarly the adequate resolution of the oscillatory basis functions). Results are reported in Figure 2. We can draw the same conclusions as in the periodic test-case: similarity of the results delivered by the mixed-order and the equal-order msHHO methods for both $k = 1$ and $k = 2$, presence of the pre-asymptotic, resonance, and asymptotic regimes, and advantages of using a higher polynomial order for the face unknowns.

To briefly assess computational costs, we compute, for those mesh levels in the pre-asymptotic or resonance regimes for which the error is minimal, the computational times to perform the offline and online steps. We report the results in Table 1. We also report the number of degrees of freedom in the global system solved in the online step. We make the experiment for the equal-order msHHO method of orders $k = 0$ and $k = 2$, for respective mesh levels $l = 2$ and $l = 1$. We do not make use of parallelism in our implementation to compute the results. Table 1 shows the interest of higher-order approximations, since a better accuracy is reached at a smaller online computational cost.

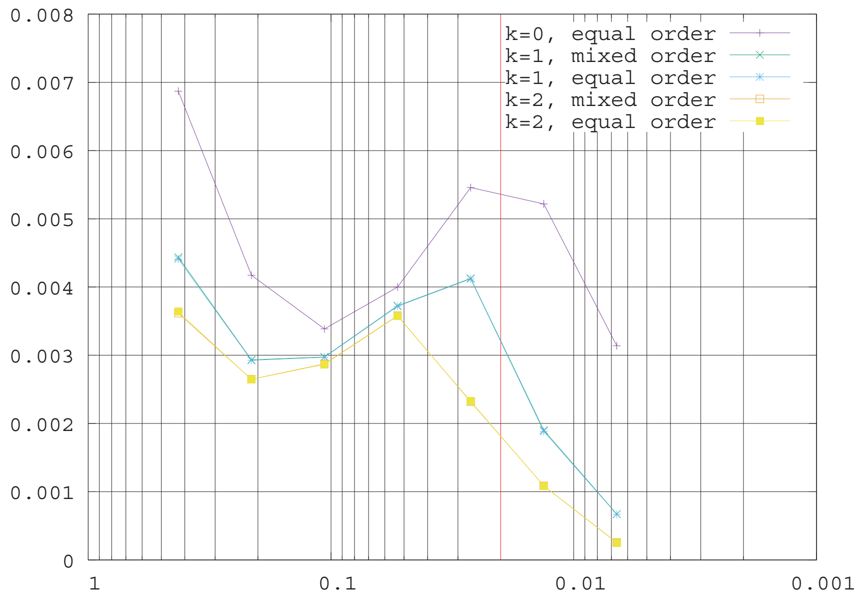

Figure 2: Locally periodic test-case: convergence results in energy-norm for mesh levels $l \in \{0, \dots, 6\}$; mixed-order msHHO method with polynomial degrees $k \in \{1, 2\}$ and equal-order msHHO method with polynomial degrees $k \in \{0, 1, 2\}$. The red vertical line indicates the value of ε.

	Energy-error	Offline time (s)	Online time (s)	#DoFs
$k = 0$ ($l = 2$)	0.00338612	254	0.026	408
$k = 2$ ($l = 1$)	0.00264648	520	0.018	288

Table 1: Offline and online computational times.

A Estimates on the First-Order Two-Scale Expansion

In this appendix, we derive various useful estimates on the first-order two-scale expansion $\mathcal{L}^1_\varepsilon(u_0)$ defined by (2.5). Except for Lemma A.4, these estimates are classical; we provide (short) proofs since we additionally track the direct dependency of the constants on the parameters α and β characterizing the spectrum of $\mathbb{A}$ and on the various length scales present in the problem.

A.1 Dual-Norm Estimates

Let D be an open, connected, polytopal subset of Ω; in this work, we will need the cases where $D = \Omega$ or where $D = T \in \mathcal{T}_H$. Let ℓ_D be a length scale associated with D, e.g., its diameter. Our goal is to bound the dual norm of the linear map such that

$$w \mapsto \mathcal{F}_\varepsilon(w) := \int_D (\mathbb{A}_\varepsilon \nabla \mathcal{L}^1_\varepsilon(u_0) - \mathbb{A}_0 \nabla u_0) \cdot \nabla w \tag{A.1}$$

for all $w \in H^1_0(D)$ (Dirichlet case), or for all $w \in H^1_\star(D) := \{w \in H^1(D) : \int_D w = 0\}$ (Neumann case); note that $\mathcal{F}_\varepsilon(w)$ does not change if the values of w are shifted by a constant.

Lemma A.1 (Dual Norm, Dirichlet Case). *Assume that the homogenized solution u_0 belongs to $H^2(D)$ and that, for any $1 \le l \le d$, the corrector μ_l belongs to $W^{1,\infty}(\mathbb{R}^d)$. Then*

$$\sup_{w \in H^1_0(D)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\nabla w\|_{L^2(D)^d}} \le c\, \beta \varepsilon |u_0|_{H^2(D)},$$

with c independent of ε, D, u_0, α, β, and possibly depending on d, and on $\max_{1 \le l \le d} \|\mu_l\|_{W^{1,\infty}(\mathbb{R}^d)}$.

Proof. For any integer $1 \le i \le d$, we have

$$\begin{aligned}
[\mathbb{A}_\varepsilon \nabla \mathcal{L}^1_\varepsilon(u_0)]_i &= \sum_{j=1}^d [\mathbb{A}_\varepsilon]_{ij} \partial_j \mathcal{L}^1_\varepsilon(u_0) \\
&= \sum_{j=1}^d [\mathbb{A}_\varepsilon]_{ij} \bigg(\partial_j u_0 + \varepsilon \sum_{l=1}^d \Big(\frac{1}{\varepsilon} \mathcal{R}_\varepsilon(\partial_j \mu_l) \partial_l u_0 + \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0 \Big) \bigg) \\
&= [\mathbb{A}_0 \nabla u_0]_i + \sum_{l=1}^d \mathcal{R}_\varepsilon(\theta^l_i) \partial_l u_0 + \varepsilon \sum_{l,j=1}^d [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0,
\end{aligned} \tag{A.2}$$

with $\theta^l_i := \mathbb{A}_{il} + \sum_{j=1}^d \mathbb{A}_{ij} \partial_j \mu_l - [\mathbb{A}_0]_{il}$ satisfying the following properties:

- $\theta^l_i \in L^\infty_{\mathrm{per}}(Q)$ by assumption on $\mathbb{A}$ and on the correctors μ_l,
- $\int_Q \theta^l_i = 0$ as a consequence of (2.2),
- $\sum_{i=1}^d \partial_i \theta^l_i = 0$ in $\mathbb{R}^d$ as a consequence of (2.3).

Adapting [37, equation (1.11)] (see also [30, Sections I.3.1 and I.3.3]), we infer that, for any integer $1 \le l \le d$, there exists a skew-symmetric matrix $\mathbb{T}^l \in W^{1,\infty}_{\mathrm{per}}(Q)^{d\times d}$, satisfying $\int_Q \mathbb{T}^l = \mathbb{0}$ and such that, for any integer $1 \le i \le d$,

$$\theta^l_i = \sum_{q=1}^d \partial_q \mathbb{T}^l_{qi}. \tag{A.3}$$

Plugging (A.3) into (A.2), we infer that, for any integer $1 \le i \le d$,

$$[\mathbb{A}_\varepsilon \nabla \mathcal{L}^1_\varepsilon(u_0)]_i - [\mathbb{A}_0 \nabla u_0]_i = \varepsilon \bigg(\sum_{l,q=1}^d \partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})) \partial_l u_0 + \sum_{l,j=1}^d [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0 \bigg).$$

Since

$$\partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})) \partial_l u_0 = \partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial_l u_0) - \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial^2_{q,l} u_0,$$

and recalling the definition (A.1) of $\mathcal{F}_\varepsilon$, this yields

$$\begin{aligned}
\mathcal{F}_\varepsilon(w) &= \varepsilon \bigg(\sum_{i,l,j=1}^d \int_D [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0 \, \partial_i w - \sum_{i,l,q=1}^d \int_D \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial^2_{q,l} u_0 \, \partial_i w \bigg) \\
&\quad + \varepsilon \sum_{i,l,q=1}^d \int_D \partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial_l u_0) \partial_i w.
\end{aligned} \tag{A.4}$$

Since $\mathbb{T}^l_{qi} = -\mathbb{T}^l_{iq}$ for any integers $1 \le i, q \le d$, we infer by integration by parts of the last term that

$$\begin{aligned}
\mathcal{F}_\varepsilon(w) &= \varepsilon \bigg(\sum_{i,l,j=1}^d \int_D [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0 \, \partial_i w - \sum_{i,l,q=1}^d \int_D \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial^2_{q,l} u_0 \, \partial_i w \bigg) \\
&\quad + \varepsilon \sum_{i,l,q=1}^d \int_{\partial D} \partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial_l u_0) n_{\partial D,i} \, w,
\end{aligned} \tag{A.5}$$

where $\boldsymbol{n}_{\partial D}$ is the unit outward normal to D. Since $w \in H^1_0(D)$, we obtain

$$\mathcal{F}_\varepsilon(w) = \varepsilon \bigg(\sum_{i,l,j=1}^d \int_D [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0 \, \partial_i w - \sum_{i,l,q=1}^d \int_D \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial^2_{q,l} u_0 \, \partial_i w \bigg).$$

Using the Cauchy–Schwarz inequality, we finally deduce that

$$\sup_{w \in H^1_0(D)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\nabla w\|_{L^2(D)^d}} \le c\,\beta\varepsilon \max_{1\le l\le d} \big(\|\mu_l\|_{L^\infty(\mathbb{R}^d)}, \beta^{-1} \|\mathbb{T}^l\|_{L^\infty(\mathbb{R}^d)^{d\times d}} \big) |u_0|_{H^2(D)}.$$

We conclude by observing that $\|\mathbb{T}^l\|_{L^\infty(\mathbb{R}^d)^{d\times d}} \le c \, \max_{1\le i\le d} \|\theta^l_i\|_{L^\infty(\mathbb{R}^d)} \le c\,\beta$. □

Lemma A.2 (Dual Norm, Neumann Case (i)). *Assume that the homogenized solution u_0 belongs to the space $H^2(D) \cap W^{1,\infty}(D)$ and that, for any $1 \le l \le d$, the corrector μ_l belongs to $W^{1,\infty}(\mathbb{R}^d)$. Then*

$$\sup_{w \in H^1_\star(D)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\nabla w\|_{L^2(D)^d}} \le c\,\beta\big(\varepsilon |u_0|_{H^2(D)} + |\partial D|^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} |u_0|_{W^{1,\infty}(D)}\big), \tag{A.6}$$

with c independent of ε, D, u_0, α, β, and possibly depending on d, and on $\max_{1\le l\le d} \|\mu_l\|_{W^{1,\infty}(\mathbb{R}^d)}$.

Proof. Our starting point is (A.4). The first two terms on the right-hand side are responsible for a contribution of order $\beta\varepsilon|u_0|_{H^2(D)}$, and it only remains to bound the last term. Following the ideas of [37, p. 29], we define, for $\eta > 0$, the domain $D_\eta := \{\boldsymbol{x} \in D : \operatorname{dist}(\boldsymbol{x}, \partial D) < \eta\}$. If η is above a critical value (which scales as ℓ_D), $D_\eta = D$, otherwise $D_\eta \subsetneq D$. We introduce the cut-off function $\zeta_\eta \in C^0(\overline{D})$ such that $\zeta_\eta \equiv 0$ on ∂D, defined by $\zeta_\eta(\boldsymbol{x}) = \frac{1}{\eta}\operatorname{dist}(\boldsymbol{x}, \partial D)$ if $\boldsymbol{x} \in D_\eta$, and $\zeta_\eta(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in D \backslash D_\eta$. We have $0 \le \zeta_\eta \le 1$ and $\max_{1\le q\le d} \|\partial_q \zeta_\eta\|_{L^\infty(D)} \le \eta^{-1}$. We first infer that

$$\varepsilon \sum_{i,l,q=1}^{d} \int_D \partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0)\partial_i w = \varepsilon \sum_{i,l,q=1}^{d} \int_{D_\eta} \partial_q((1-\zeta_\eta)\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0)\partial_i w,$$

since $(1 - \zeta_\eta)$ vanishes identically on $D \setminus D_\eta$ and since

$$\sum_{i,l,q=1}^{d} \int_D \partial_q(\zeta_\eta \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0)\partial_i w = 0$$

as can be seen by integration by parts, using the fact that $\mathbb{T}^l_{qi} = -\mathbb{T}^l_{iq}$ for any integers $1 \le i, q \le d$, and the fact that ζ_η vanishes identically on ∂D. Then, accounting for the fact that

$$\varepsilon\,\partial_q((1-\zeta_\eta)\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0) = -\varepsilon\,\partial_q\zeta_\eta\,\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0 + (1-\zeta_\eta)\mathcal{R}_\varepsilon(\partial_q \mathbb{T}^l_{qi})\partial_l u_0 + \varepsilon(1-\zeta_\eta)\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial^2_{q,l} u_0,$$

we infer that

$$\begin{aligned}\Big|\varepsilon \sum_{i,l,q=1}^{d} \int_D \partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0)\partial_i w\Big| \le c\Big[&|D_\eta|^{\frac{1}{2}}\Big(\frac{\varepsilon}{\eta} + 1\Big)\Big(\max_{1\le l\le d} \|\mathbb{T}^l\|_{W^{1,\infty}(\mathbb{R}^d)^{d\times d}}\Big)|u_0|_{W^{1,\infty}(D)} \\ &+ \varepsilon\Big(\max_{1\le l\le d} \|\mathbb{T}^l\|_{L^\infty(\mathbb{R}^d)^{d\times d}}\Big)|u_0|_{H^2(D)}\Big]\|\nabla w\|_{L^2(D)^d}.\end{aligned}$$

Using now the estimate $|D_\eta| \le \eta|\partial D|$, the fact that $\max_{1\le l\le d} \|\mathbb{T}^l\|_{W^{1,\infty}(\mathbb{R}^d)^{d\times d}} \le c\,\beta$, and since the function $\eta \mapsto \frac{\varepsilon}{\sqrt{\eta}} + \sqrt{\eta}$ is minimal for $\eta = \varepsilon$, we finally infer the bound (A.6). ☐

Remark A.3 (Weaker Regularity Assumption). Without the regularity assumption $u_0 \in W^{1,\infty}(D)$, one can still invoke a Sobolev embedding since $u_0 \in H^2(D)$. The second term between the parentheses on the right-hand side of (A.6) becomes

$$c(p)(|\partial D|\varepsilon\ell_D^{-d})^{\frac{1}{2}-\frac{1}{p}}(|u_0|_{H^1(D)} + \ell_D|u_0|_{H^2(D)}),$$

where $p = 6$ for $d = 3$ and p can be taken as large as wanted for $d = 2$ (note that $c(p) \to +\infty$ when $p \to +\infty$ in that case). The derivation of estimates in this setting is considered in [43]. Therein, the authors claim that their method is able to get rid of the resonance error (without oversampling). We believe there is an issue with the bound [43, equation (27)], which should exhibit the resonant contribution $(\frac{\varepsilon}{\ell_D})^{\frac{1}{2}-\frac{1}{p}}|u_0|_{H^1(D)}$.

Lemma A.4 (Dual Norm, Neumann Case (ii)). *Assume that $D = T \in \mathcal{T}_H$, where $\mathcal{T}_H$ is a member of an admissible mesh sequence in the sense of Definition 3.1; set $\ell_D = H_T$. Assume that the homogenized solution u_0 belongs to $H^3(D)$ and that there is $\kappa > 0$ so that $\mathbb{A} \in C^{0,\kappa}(\mathbb{R}^d; \mathbb{R}^{d\times d})$. Then*

$$\sup_{w \in H^1_\star(D)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\nabla w\|_{L^2(D)^d}} \le c\,\beta\Big((\varepsilon + (\varepsilon\ell_D)^{\frac{1}{2}})|u_0|_{H^2(D)} + \varepsilon\ell_D|u_0|_{H^3(D)} + \varepsilon^{\frac{1}{2}}\ell_D^{-\frac{1}{2}}|u_0|_{H^1(D)}\Big),$$

with c independent of ε, D, u_0, α, β, and possibly depending on d, γ, and $\|\frac{\mathbb{A}}{\beta}\|_{C^{0,\kappa}(\mathbb{R}^d;\mathbb{R}^{d\times d})}$.

Proof. We proceed as in the proof of Lemma A.1. Concerning the regularity of θ^l_i, we now have $\theta^l_i \in C^{0,\iota}(\mathbb{R}^d)$ for some $\iota > 0$ since the Hölder continuity of $\mathbb{A}$ on $\mathbb{R}^d$ implies the Hölder continuity of μ_l and $\nabla\mu_l$ on $\mathbb{R}^d$ for any $1 \le l \le d$; cf., e.g., [29, Theorem 8.22 and Corollary 8.36]. Following [37, pp. 6–7] and [39, pp. 131–132], we infer that the skew-symmetric matrix $\mathbb{T}^l$ is such that $\mathbb{T}^l \in C^1(\mathbb{R}^d)^{d\times d}$. Our starting point is (A.5). The first two terms on the right-hand side are responsible for a contribution of order $\beta\varepsilon|u_0|_{H^2(D)}$, and it only remains to bound the last term. We have

$$\begin{aligned}\varepsilon \sum_{i,l,q=1}^{d} \int_{\partial D} \partial_q(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0) n_{\partial D,i}\, w &= \varepsilon \sum_{i,l,q=1}^{d} \int_{\partial D} \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial^2_{q,l} u_0\, n_{\partial D,i}\, w + \sum_{i,l,q=1}^{d} \int_{\partial D} \mathcal{R}_\varepsilon(\partial_q \mathbb{T}^l_{qi})\partial_l u_0\, n_{\partial D,i}\, w \\ &=: \mathfrak{T}_1 + \mathfrak{T}_2.\end{aligned}$$

Using the Cauchy–Schwarz inequality and the trace inequality (3.2), the first term on the right-hand side can be estimated as

$$|\mathfrak{T}_1| \le c\,\beta\varepsilon\ell_D^{-1}(|u_0|_{H^2(D)} + \ell_D|u_0|_{H^3(D)})(\|w\|_{L^2(D)} + \ell_D\|\nabla w\|_{L^2(D)^d}),$$

since $\max_{1\le l\le d}\|\mathbb{T}^l\|_{C^0(\mathbb{R}^d)^{d\times d}} \le c\,\beta$. Observing that $\int_D w = 0$, we can use the Poincaré inequality (3.3) to infer that

$$|\mathfrak{T}_1| \le c\,\beta\varepsilon(|u_0|_{H^2(D)} + \ell_D|u_0|_{H^3(D)})\|\nabla w\|_{L^2(D)^d}.$$

To estimate the second term on the right-hand side, we adapt the ideas from [39, Lemma 4.6]. Considering the matching simplicial sub-mesh of D, let us collect in the set $\mathfrak{F}_D$ all the sub-faces composing the boundary of D. Then we can write

$$\mathfrak{T}_2 = \sum_{\sigma\in\mathfrak{F}_D}\sum_{l=1}^{d}\sum_{q=1}^{d}\sum_{q<i\le d}\int_\sigma \mathcal{R}_\varepsilon(\nabla\mathbb{T}^l_{qi})\cdot\boldsymbol{\tau}^{qi}_\sigma\,\partial_l u_0\, w,$$

where the vectors $\boldsymbol{\tau}^{qi}_\sigma$ are such that $\|\boldsymbol{\tau}^{qi}_\sigma\|_{\ell^2} \le 1$ and $\boldsymbol{\tau}^{qi}_\sigma\cdot\boldsymbol{n}_{\partial D|\sigma} = 0$. Then using a straightforward adaptation of the result in [39, Lemma 4.6], and since $\max_{1\le l\le d}\|\mathbb{T}^l\|_{C^1(\mathbb{R}^d)^{d\times d}} \le c\,\beta$, we infer that

$$\left|\int_\sigma \mathcal{R}_\varepsilon(\nabla\mathbb{T}^l_{qi})\cdot\boldsymbol{\tau}^{qi}_\sigma\,\partial_l u_0\, w\right| \le c\,\beta\varepsilon^{\frac12}H_S^{-\frac32}(|u_0|_{H^1(S)} + H_S|u_0|_{H^2(S)})(\|w\|_{L^2(S)} + H_S\|\nabla w\|_{L^2(S)^d}),$$

where S is the simplicial sub-cell of D having σ as face. Collecting the contributions of all the sub-faces $\sigma\in\mathfrak{F}_D$ and using the mesh regularity assumptions on D, we infer that

$$|\mathfrak{T}_2| \le c\,\beta\varepsilon^{\frac12}\ell_D^{-\frac32}(|u_0|_{H^1(D)} + \ell_D|u_0|_{H^2(D)})(\|w\|_{L^2(D)} + \ell_D\|\nabla w\|_{L^2(D)^d}).$$

Finally, invoking the Poincaré inequality (3.3) since w has zero mean-value in D yields

$$|\mathfrak{T}_2| \le c\,\beta\varepsilon^{\frac12}\ell_D^{-\frac12}(|u_0|_{H^1(D)} + \ell_D|u_0|_{H^2(D)})\|\nabla w\|_{L^2(D)^d}.$$

Collecting the above bounds on $\mathfrak{T}_1$ and $\mathfrak{T}_2$ concludes the proof. □

A.2 Global Energy-Norm Estimate

Lemma A.5 (Energy-Norm Estimate). *Assume that the homogenized solution u_0 belongs to $H^2(\Omega)\cap W^{1,\infty}(\Omega)$, and that, for any $1\le l\le d$, the corrector μ_l belongs to $W^{1,\infty}(\mathbb{R}^d)$. Then*

$$\|\mathbb{A}_\varepsilon^{\frac12}\nabla(u_\varepsilon - \mathcal{L}^1_\varepsilon(u_0))\|_{L^2(\Omega)^d} \le c\,\beta^{\frac12}(\rho^{\frac12}\varepsilon\,|u_0|_{H^2(\Omega)} + |\partial\Omega|^{\frac12}\varepsilon^{\frac12}\,|u_0|_{W^{1,\infty}(\Omega)}),$$

with c independent of ε, Ω, u_0, α, β, and possibly depending on d, and on $\max_{1\le l\le d}\|\mu_l\|_{W^{1,\infty}(\mathbb{R}^d)}$.

Proof. The regularity assumptions on u_0 and the correctors imply $(u_\varepsilon - \mathcal{L}^1_\varepsilon(u_0)) \in H^1(\Omega)$; however, we do not have $(u_\varepsilon - \mathcal{L}^1_\varepsilon(u_0)) \in H^1_0(\Omega)$. Following the ideas in [37, p. 28], we define, for $\eta > 0$, the domain $\Omega_\eta := \{\boldsymbol{x}\in\Omega : \mathrm{dist}(\boldsymbol{x},\partial\Omega) < \eta\}$. If η is above a critical value, $\Omega_\eta = \Omega$, otherwise $\Omega_\eta \subsetneq \Omega$. We introduce the cut-off function $\zeta_\eta \in C^0(\overline{\Omega})$ such that $\zeta_\eta \equiv 0$ on $\partial\Omega$, defined by $\zeta_\eta(\boldsymbol{x}) = \frac{1}{\eta}\mathrm{dist}(\boldsymbol{x},\partial\Omega)$ if $\boldsymbol{x}\in\Omega_\eta$, and $\zeta_\eta(\boldsymbol{x}) = 1$ if $\boldsymbol{x}\in\Omega\setminus\Omega_\eta$. We have $0\le\zeta_\eta\le 1$ and $\max_{1\le i\le d}\|\partial_i\zeta_\eta\|_{L^\infty(\Omega)} \le \eta^{-1}$. The function ζ_η allows us to define a corrected first-order two-scale expansion $\mathcal{L}_\varepsilon^{1,0}(u_0) := u_0 + \varepsilon\zeta_\eta \sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l)\partial_l u_0$ such that $(u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0)) \in H_0^1(\Omega)$. We start with the triangle inequality:

$$\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0))\|_{L^2(\Omega)^d} \le \|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0))\|_{L^2(\Omega)^d} + \|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0))\|_{L^2(\Omega)^d}. \tag{A.7}$$

Let us focus on the first term on the right-hand side of (A.7). We have

$$\begin{aligned}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0))\|^2_{L^2(\Omega)^d} &= \int_\Omega \mathbb{A}_\varepsilon\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0))\cdot\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0)) \\ &\quad + \int_\Omega \mathbb{A}_\varepsilon\nabla(\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0))\cdot\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0)).\end{aligned}$$

Since $(u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0)) \in H_0^1(\Omega)$, we infer that

$$\begin{aligned}\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0))\|_{L^2(\Omega)^d} &\le \alpha^{-\frac{1}{2}} \sup_{w\in H_0^1(\Omega)} \frac{|\int_\Omega \mathbb{A}_\varepsilon\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0))\cdot\nabla w|}{\|\nabla w\|_{L^2(\Omega)^d}} \\ &\quad + \|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0))\|_{L^2(\Omega)^d}.\end{aligned} \tag{A.8}$$

Since $\int_\Omega \mathbb{A}_\varepsilon\nabla u_\varepsilon\cdot\nabla w = \int_\Omega \mathbb{A}_0\nabla u_0\cdot\nabla w$ for any $w \in H_0^1(\Omega)$ in view of (1.1) and (2.4), estimates (A.7) and (A.8) lead to

$$\|\mathbb{A}_\varepsilon^{\frac{1}{2}}\nabla(u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0))\|_{L^2(\Omega)^d} \le \alpha^{-\frac{1}{2}} \sup_{w\in H_0^1(\Omega)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\nabla w\|_{L^2(\Omega)^d}} + 2\beta^{\frac{1}{2}}\|\nabla(\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0))\|_{L^2(\Omega)^d}, \tag{A.9}$$

recalling that $\mathcal{F}_\varepsilon(w) = \int_\Omega(\mathbb{A}_\varepsilon\nabla\mathcal{L}_\varepsilon^1(u_0) - \mathbb{A}_0\nabla u_0)\cdot\nabla w$. Since we can bound the first term on the right-hand side of (A.9) using Lemma A.1 (with $D = \Omega$), it remains to estimate the second term. Owing to the definition of ζ_η, we infer that

$$\|\nabla(\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0))\|_{L^2(\Omega)^d} = \varepsilon\Big\|\nabla\Big((1-\zeta_\eta)\sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l)\partial_l u_0\Big)\Big\|_{L^2(\Omega_\eta)^d}.$$

For any integer $1 \le i \le d$, we have

$$\partial_i\Big((1-\zeta_\eta)\sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l)\partial_l u_0\Big) = -\partial_i\zeta_\eta\sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l)\partial_l u_0 + \frac{(1-\zeta_\eta)}{\varepsilon}\sum_{l=1}^d \mathcal{R}_\varepsilon(\partial_i\mu_l)\partial_l u_0 + (1-\zeta_\eta)\sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l)\partial^2_{i,l}u_0,$$

and using the properties of the cut-off function ζ_η, we infer that

$$\varepsilon\Big\|\nabla\Big((1-\zeta_\eta)\sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l)\partial_l u_0\Big)\Big\|_{L^2(\Omega_\eta)^d} \le c\Big(|\Omega_\eta|^{\frac{1}{2}}\Big(\frac{\varepsilon}{\eta}+1\Big)|u_0|_{W^{1,\infty}(\Omega)} + \varepsilon|u_0|_{H^2(\Omega)}\Big).$$

Since $|\Omega_\eta| \le |\partial\Omega|\eta$, and choosing $\eta = \varepsilon$ to minimize the function $\eta \mapsto \frac{\varepsilon}{\sqrt{\eta}} + \sqrt{\eta}$, we can conclude the proof (note that $\rho \ge 1$ by definition). □

Acknowledgment: The authors are thankful to Alexei Lozinski (LMB, Université de Franche-Comté) for fruitful discussions on the topic.

References

[1] A. Abdulle, W. E, B. Engquist and E. Vanden-Eijnden, The heterogeneous multiscale method, *Acta Numer.* **21** (2012), 1–87.
[2] G. Allaire, *Shape Optimization by the Homogenization Method*, Appl. Math. Sci. 146, Springer, New York, 2002.
[3] G. Allaire and R. Brizzi, A multiscale finite element method for numerical homogenization, *SIAM Multiscale Model. Simul.* **4** (2005), no. 3, 790–812.
[4] R. Araya, C. Harder, D. Paredes and F. Valentin, Multiscale hybrid-mixed method, *SIAM J. Numer. Anal.* **51** (2013), no. 6, 3505–3531.

[5] D. N. Arnold, F. Brezzi, B. Cockburn and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *SIAM J. Numer. Anal.* **39** (2002), no. 5, 1749–1779.
[6] B. Ayuso de Dios, K. Lipnikov and G. Manzini, The nonconforming virtual element method, *ESAIM Math. Model. Numer. Anal. (M2AN)* **50** (2016), no. 3, 879–904.
[7] F. Bassi, L. Botti, A. Colombo, D. A. Di Pietro and P. Tesini, On the flexibility of agglomeration-based physical space discontinuous Galerkin discretizations, *J. Comput. Phys.* **231** (2012), no. 1, 45–65.
[8] M. Bebendorf, A note on the Poincaré inequality for convex domains, *Z. Anal. Anwend.* **22** (2003), no. 4, 751–756.
[9] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini and A. Russo, Basic principles of virtual element methods, *Math. Models Meth. Appl. Sci. (M3AS)* **23** (2013), 199–214.
[10] A. Cangiani, E. H. Georgoulis and P. Houston, *hp*-version discontinuous Galerkin methods on polygonal and polyhedral meshes, *Math. Models Meth. Appl. Sci. (M3AS)* **24** (2014), no. 10, 2009–2041.
[11] E. T. Chung, S. Fu and Y. Yang, An enriched multiscale mortar space for high contrast flow problems, *Commun. Comput. Phys.* **23** (2018), no. 2, 476–499.
[12] M. Cicuttin, D. A. Di Pietro and A. Ern, Implementation of discontinuous skeletal methods on arbitrary-dimensional, polytopal meshes using generic programming, *J. Comput. Appl. Math.* (2017), DOI 10.1016/j.cam.2017.09.017.
[13] B. Cockburn, Static condensation, hybridization, and the devising of the HDG methods, in: *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, Lect. Notes Comput. Sci. Eng. 114, Springer, Cham (2016), 129–177.
[14] B. Cockburn, D. A. Di Pietro and A. Ern, Bridging the hybrid high-order and hybridizable discontinuous Galerkin methods, *ESAIM Math. Model. Numer. Anal. (M2AN)* **50** (2016), no. 3, 635–650.
[15] B. Cockburn, J. Gopalakrishnan and R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second-order elliptic problems, *SIAM J. Numer. Anal.* **47** (2009), no. 2, 1319–1365.
[16] D. A. Di Pietro and A. Ern, *Mathematical Aspects of Discontinuous Galerkin Methods*, Math. Appl. 69, Springer, Berlin, 2012.
[17] D. A. Di Pietro and A. Ern, A hybrid high-order locking-free method for linear elasticity on general meshes, *Comput. Methods Appl. Mech. Engrg.* **283** (2015), 1–21.
[18] D. A. Di Pietro and A. Ern, Hybrid high-order methods for variable-diffusion problems on general meshes, *C. R. Acad. Sci. Paris Ser. I* **353** (2015), 31–34.
[19] D. A. Di Pietro, A. Ern and S. Lemaire, An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators, *Comput. Methods Appl. Math.* **14** (2014), no. 4, 461–472.
[20] D. A. Di Pietro, A. Ern and S. Lemaire, A review of hybrid high-order methods: Formulations, computational aspects, comparison with other methods, in: *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations* Lect. Notes Comput. Sci. Eng. 114, Springer, Cham (2016), 205–236.
[21] W. E and B. Engquist, The heterogeneous multiscale methods, *Commun. Math. Sci.* **1** (2003), 87–132.
[22] Y. Efendiev, J. Galvis and T. Y. Hou, Generalized multiscale finite element methods (GMsFEM), *J. Comput. Phys.* **251** (2013), 116–135.
[23] Y. Efendiev and T. Y. Hou, *Multiscale Finite Element Methods – Theory and Applications*, Surv. Tutor. Appl. Math. Sci. 4, Springer, New York, 2009.
[24] Y. Efendiev, T. Y. Hou and X.-H. Wu, Convergence of a nonconforming multiscale finite element method, *SIAM J. Numer. Anal.* **37** (2000), no. 3, 888–910.
[25] Y. Efendiev, R. Lazarov, M. Moon and K. Shi, A spectral multiscale hybridizable discontinuous Galerkin method for second order elliptic problems, *Comput. Methods Appl. Mech. Engrg.* **292** (2015), 243–256.
[26] Y. Efendiev, R. Lazarov and K. Shi, A multiscale HDG method for second order elliptic equations. Part I. Polynomial and homogenization-based multiscale spaces, *SIAM J. Numer. Anal.* **53** (2015), no. 1, 342–369.
[27] A. Ern and J.-L. Guermond, Finite element quasi-interpolation and best approximation, *ESAIM Math. Model. Numer. Anal. (M2AN)* **51** (2017), no. 4, 1367–1385.
[28] A. Ern and M. Vohralík, Stable broken H^1 and $H(\mathrm{div})$ polynomial extensions for polynomial-degree-robust potential and flux reconstruction in three space dimensions, preprint (2016), https://hal.inria.fr/hal-01422204.
[29] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Classics Math., Springer, Berlin, 2001.
[30] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer, Berlin, 1986.
[31] A. Gloria, Numerical homogenization: Survey, new results, and perspectives, *ESAIM Proc.* **37** (2012), 50–116.
[32] P. Henning and D. Peterseim, Oversampling for the multiscale finite element method, *SIAM Multiscale Model. Simul.* **11** (2013), no. 4, 1149–1175.
[33] J. S. Hesthaven, S. Zhang and X. Zhu, High-order multiscale finite element method for elliptic problems, *SIAM Multiscale Model. Simul.* **12** (2014), no. 2, 650–666.
[34] T. Y. Hou and X.-H. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, *J. Comput. Phys.* **134** (1997), 169–189.

[35] T. Y. Hou, X.-H. Wu and Z. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients, *Math. Comp.* **68** (1999), no. 227, 913–943.
[36] T. Y. Hou, X.-H. Wu and Y. Zhang, Removing the cell resonance error in the multiscale finite element method via a Petrov–Galerkin formulation, *Commun. Math. Sci.* **2** (2004), no. 2, 185–205.
[37] V. V. Jikov, S. M. Kozlov and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer, Berlin, 1994.
[38] A. Konaté, *Méthode multi-échelle pour la simulation d'écoulements miscibles en milieux poreux*, Ph.D. thesis, Université Pierre et Marie Curie, 2017, https://tel.archives-ouvertes.fr/tel-01558994.
[39] C. Le Bris, F. Legoll and A. Lozinski, MsFEM à la Crouzeix–Raviart for highly oscillatory elliptic problems, *Chin. Ann. Math. Ser. B* **34** (2013), no. 1, 113–138.
[40] C. Le Bris, F. Legoll and A. Lozinski, An MsFEM-type approach for perforated domains, *SIAM Multiscale Model. Simul.* **12** (2014), no. 3, 1046–1077.
[41] A. Målqvist and D. Peterseim, Localization of elliptic multiscale problems, *Math. Comp.* **83** (2014), 2583–2603.
[42] L. Mu, J. Wang and X. Ye, A weak Galerkin generalized multiscale finite element method, *J. Comput. Appl. Math.* **305** (2016), 68–81.
[43] D. Paredes, F. Valentin and H. M. Versieux, On the robustness of multiscale hybrid-mixed methods, *Math. Comp.* **86** (2017), 525–548.
[44] N. Sukumar and A. Tabarraei, Conforming polygonal finite elements, *Internat. J. Numer. Methods Engrg.* **61** (2004), no. 12, 2045–2066.
[45] A. Veeser and R. Verfürth, Poincaré constants for finite element stars, *IMA J. Numer. Anal.* **32** (2012), no. 1, 30–47.
[46] E. L. Wachspress, *A Rational Finite Element Basis*, Math. Sci. Eng. 114, Academic Press, New York, 1975.
[47] J. Wang and X. Ye, A weak Galerkin finite element method for second-order elliptic problems, *J. Comput. Appl. Math.* **241** (2013), 103–115.