A Multi-scale Hybrid High-Order method

Matteo Cicuttin, Alexandre Ern, Simon Lemaire

A Multi-scale Hybrid High-Order method

Matteo Cicuttin*[1], Alexandre Ern†[1], and Simon Lemaire‡[2]

[1]Université Paris-Est, CERMICS (ENPC), 6–8 avenue Blaise Pascal, 77455 Marne-la-Vallée CEDEX 2, France
[2]École Polytechnique Fédérale de Lausanne (EPFL), SB-MATHICSE-ANMC, Station 8, 1015 Lausanne, Switzerland

February 14, 2017

Abstract

We devise a multi-scale Hybrid High-Order (HHO) method. The method hinges on (hybrid) discrete unknowns that are polynomials attached to mesh elements and faces, and on a multi-scale reconstruction operator, that maps onto a fine-scale space spanned by oscillatory basis functions. The method handles arbitrary orders of approximation $k \geqslant 0$, and is applicable on general meshes. For face-based unknowns that are polynomials of degree k, we devise two versions of the method, depending on the polynomial degree $(k-1)$ or k of cell-based unknowns. We prove, in the case of periodic coefficients, an energy-error estimate of the form $\left(\varepsilon^{1/2} + H^{k+1} + \varepsilon^{1/2}H^{-1/2}\right)$.

1 Introduction

Over the last few years, a great deal of effort has been devoted to the design of new-generation arbitrary-order polytopal discretization methods. Such methods are approaches that are capable of handling meshes with polytopal cells of (almost) arbitrary shapes. Classical approaches encompass the (polytopal) Finite Element (FE) [40, 38], and the Discontinuous Galerkin (DG) [5, 15, 9] methods. Classical methods however suffer from some drawbacks: for the FE method, the difficulty to construct basis functions (due to continuity requirements) and the fact that they are usually non-polynomial, and for the DG method, the rapidly increasing (with respect to the order of the method) number of globally coupled degrees of freedom.

More recently, a new paradigm has emerged. The main idea is to consider, locally in each cell, a discrete function space that encompasses all the functions that are solution

*matteo.cicuttin@enpc.fr
†alexandre.ern@enpc.fr
‡Corresponding author: simon.lemaire@epfl.ch

of the equation under study, supplemented by polynomial right-hand side and polynomial (Dirichlet or Neumann) boundary conditions of given degrees. The discrete unknowns are thus polynomials attached to the cells and to the faces of the mesh. The global space can be either continuous along the skeleton, or discontinuous. The specificity of new-generation polytopal discretization methods then comes from the fact that one can only keep the functions from the discrete space that are sufficient to give optimal approximation properties to the method (typically, polynomial functions of one degree higher in the cell). At the end of the day, the basis functions that enter the computations are all polynomial, and the non-polynomial ones are handled in a finely tuned stabilization term. These methods can be referred to as skeletal, since cell-based discrete unknowns can always be locally eliminated by static condensation, hence leading to global systems posed in terms of skeletal unknowns only. This obviously reduces (compared, e.g., to DG methods) the dependency with respect to the order of the method of the number of globally coupled degrees of freedom. A globally conforming example of a new-generation polytopal discretization method is the Virtual Element (VE) [8] method, whereas globally non-conforming examples include the Hybridizable Discontinuous Galerkin (HDG) [14] method, the related Weak Galerkin (WG) [41] method (proved equivalent to HDG in [12]), and the Hybrid High-Order (HHO) [17] method, that has been bridged to HDG in [13] (the latter reference also fits into the HHO framework the non-conforming VE method of [6], up to equivalent stabilization).

The focus here is on HHO methods. These methods offer several assets, like, e.g., a dimension-independent construction, and local conservativity. We are interested in diffusion problems featuring heterogeneous/anisotropic coefficients. The case of mildly heterogeneous (i.e., slowly varying) coefficients has already been treated in [16] (see also [18]), where error estimates tracking the dependency of the approximation with respect to the local heterogeneity/anisotropy ratios have been derived. In this article, we are interested in highly oscillatory problems. Let Ω be an open, bounded, connected polytopal subset of $\mathbb{R}^d$, $d \in \{2,3\}$. Let $\varepsilon > 0$, supposedly much smaller than the diameter of the domain Ω. We consider the problem

$$\begin{cases} -\operatorname{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} u_\varepsilon) = f & \text{in } \Omega, \\ u_\varepsilon = 0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

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. The parameter ε is meant to encode the fine-scale oscillations of the coefficients. It is well-known that the H^{k+2}-norm of the solution u_ε to Problem (1) scales as $\varepsilon^{-(k+1)}$, meaning that mono-scale methods (including the mono-scale HHO method of order $k \geqslant 0$ of [16]) provide an energy-norm decay of the error of order $(h/\varepsilon)^{k+1}$. 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 (think, e.g., of a time-dependent model), a mono-scale solve is hence unaffordable. In that context, multi-scale methods may be preferred. Multi-scale 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 an approximation on a coarse mesh with size $H \gg \varepsilon$, using oscillatory basis functions computed in the offline step).

Multi-scale approximation methods on classical element shapes (such as simplices or quadrangles/hexahedra) have been extensively studied in the literature. Examples include, e.g., the Multi-scale Finite Element (MsFE) [28, 29, 20] method (with lowest-order error bound of the form $\left(\varepsilon^{1/2} + H + \varepsilon^{1/2}H^{-1/2}\right)$ in the periodic case), its variant using oversampling [28, 21] (with improved error bound of the form $\left(\varepsilon^{1/2} + H + \varepsilon H^{-1}\right)$ in the periodic case), the Petrov-Galerkin variant of MsFE using oversampling [30], and more recently, the MsFE method à la Crouzeix–Raviart of [33, 34] (with upper bound of the form $\left(\varepsilon^{1/2} + H + \varepsilon^{1/2}H^{-1/2}\right)$ in the periodic case). This list is far from being exhaustive. Present research directions mainly focus on reducing the cell resonance error by proposing adequate local decompositions (see, e.g., [35, 32]). Note that there exist also different paradigms to approximate oscillatory problems, like the Heterogeneous Multi-scale Method (HMM) [19, 1], whose focus is more on computing an approximation of the homogenized solution instead of computing the oscillatory one; in that sense, HMM is more a numerical homogenization approach. Back to multi-scale methods, attempts to design multi-scale (arbitrary-order) *polytopal* methods include the work of Efendiev et al. [23, 22] in the HDG context (see also [10], and [36] in the WG context), and the work of Paredes, Valentin and Versieux [37] in the context of Multi-scale Hybrid-Mixed (MHM) [4] methods.

In this work, we devise a multi-scale HHO (MsHHO) method, which can be seen as a generalization to arbitrary order and general element shapes of the MsFE method à la Crouzeix–Raviart of Le Bris, Legoll and Lozinski [33, 34]. Thus, our goal is to propose and analyze (under the classical assumption of periodic coefficients) a multi-scale *arbitrary-order and polytopal* method, using the quite general framework of HHO methods. Two MsHHO methods are proposed. Both employ polynomials of order $k \geqslant 0$ for the face-based unknowns, whereas the cell-based unknowns can be polynomials of order $(k-1)$ (if $k \geqslant 1$) or k. We prove for both methods an energy-error estimate of the form $\left(\varepsilon^{1/2} + H^{k+1} + \varepsilon^{1/2}H^{-1/2}\right)$ in the periodic case. To motivate the use of a high-order method, we note that this upper bound, say $f_k(H)$, is minimal for $H_k = \left(\varepsilon^{1/2}/2(k+1)\right)^{2/(2k+3)}$, and as $k \geqslant 0$ increases, H_k increases while $f_k(H_k)$ decreases. We also track in the error bounds the dependency upon the global heterogeneity/anisotropy ratio, exhibiting a dependency that is reminiscent of the mono-scale HHO method of [16] with piecewise non-constant diffusivity. The error estimates we derive are sharper (in the sense that they describe all the regimes observed in practice) than the one derived in [37] in the context of MHM methods. Our fine-scale space construction is close to the (polynomial-based) one advocated in [23] in the HDG context. However, the two methods differ, both in the construction and in the analysis (in the latter reference, the analysis is sharp only for $H \ll \varepsilon$).

The article is organized as follows. In Sections 2 and 3 we introduce, respectively, the continuous and discrete settings. In particular, we define the notion of admissible mesh sequence. 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. We also detail the offline/online organization of the computations. 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 $1 \leqslant p \leqslant +\infty$ and $m \in \mathbb{N}^\star$, the following periodic spaces:

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

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 $0 < a \leqslant b$,

$$\mathcal{S}^b_a := \left\{ \mathbb{M} \in \mathcal{S}_d(\mathbb{R}) \mid \forall \boldsymbol{\xi} \in \mathbb{R}^d,\ a|\boldsymbol{\xi}|^2 \leqslant \mathbb{M}\boldsymbol{\xi}{\cdot}\boldsymbol{\xi} \leqslant b|\boldsymbol{\xi}|^2 \right\}.$$

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

$$\mathbb{A}(\cdot) \in \mathcal{S}^\beta_\alpha \text{ a.e. in } \mathbb{R}^d. \tag{2}$$

Assumption (2) ensures that $\mathbb{A}_\varepsilon \in L^\infty(\Omega; \mathbb{R}^{d\times d})$ is such that $\mathbb{A}_\varepsilon(\cdot) \in \mathcal{S}^\beta_\alpha$ a.e. in Ω for any $\varepsilon > 0$, and hence guarantees the existence and uniqueness of the solution to (1) in $H^1_0(\Omega)$ for any $\varepsilon > 0$. More importantly, the assumption (2) ensures that the (whole) family $(\mathbb{A}_\varepsilon)_{\varepsilon>0}$ G-converges [3, Section 1.3.2] to some constant symmetric matrix $\mathbb{A}_0 \in \mathcal{S}^\beta_\alpha$. Henceforth, we denote $\rho := \beta/\alpha \geqslant 1$ 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 $1 \leqslant i,j \leqslant d$,

$$[\mathbb{A}_0]_{ij} = \int_Q \mathbb{A}\,(\boldsymbol{e}_j + \boldsymbol{\nabla}\mu_j){\cdot}(\boldsymbol{e}_i + \boldsymbol{\nabla}\mu_i) = \int_Q \mathbb{A}\,(\boldsymbol{e}_j + \boldsymbol{\nabla}\mu_j){\cdot}\boldsymbol{e}_i, \tag{3}$$

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

$$\begin{cases} -\operatorname{div}(\mathbb{A}(\boldsymbol{\nabla}\mu_l + \boldsymbol{e}_l)) = 0 & \text{in } \mathbb{R}^d, \\ \mu_l \text{ is } \mathbb{Z}^d\text{-periodic.} \end{cases} \tag{4}$$

For further use, we also define the linear operator $\mathcal{R}_\varepsilon : L^p_{\mathrm{per}}(Q) \to L^p(\Omega)$, $1 \leqslant p \leqslant +\infty$, 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 $1 \leqslant i,j \leqslant d$, 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 $1 \leqslant l \leqslant d$.

The homogenized problem reads

$$\begin{cases} -\mathrm{div}(\mathbb{A}_0\boldsymbol{\nabla}u_0) = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega. \end{cases} \tag{5}$$

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

$$\mathcal{L}^1_\varepsilon(u_0) := u_0 + \varepsilon\sum_{l=1}^{d}\mathcal{R}_\varepsilon(\mu_l)\partial_l u_0. \tag{6}$$

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 $H \in \mathcal{H}$, 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 $\mathcal{F}^{\mathrm{i}}_H$, boundary faces in $\mathcal{F}^{\mathrm{b}}_H$, and we let $\mathcal{F}_H := \mathcal{F}^{\mathrm{i}}_H \cup \mathcal{F}^{\mathrm{b}}_H$. The diameter of a face $F \in \mathcal{F}_H$ is denoted H_F. For all $T \in \mathcal{T}_H$, we define $\mathcal{F}_T := \{F \in \mathcal{F}_H \mid F \subset \partial T\}$ the set of faces lying on 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. [15, Section 1.4] and [18, 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 $\mathcal{F}_H$) such that 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, $\gamma H_S \leqslant R_S$;

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

Two classical consequences of Definition 3.1 are that, for any mesh $\mathcal{T}_H$ belonging to an admissible mesh sequence, (i) the quantity $\mathrm{card}(\mathcal{F}_T)$ is bounded independently of the diameter H_T for all $T \in \mathcal{T}_H$ [15, Lemma 1.41], and (ii) mesh faces have a comparable diameter to the diameter of the cells to which they belong [15, Lemma 1.42].

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

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

Let a mesh $\mathcal{T}_H$ be given. For any $T \in \mathcal{T}_H$, $\mathbb{P}^q_d(T)$ is composed of the restriction to T of polynomials in $\mathbb{P}^q_d$, and for any $F \in \mathcal{F}_H$, $\mathbb{P}^q_{d-1}(F)$ is composed of the restriction to F of polynomials in $\mathbb{P}^q_d$ (this space can also be described as the restriction to F of polynomials in $\mathbb{P}^q_{d-1} \circ \Theta^{-1}$, 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:

$$\mathbb{P}^q_{d-1}(\mathcal{F}_T) := \left\{ v \in L^2(\partial T) \mid v_{|F} \in \mathbb{P}^q_{d-1}(F)\ \forall F \in \mathcal{F}_T \right\}.$$

The term 'broken' refers to the fact that no continuity is required between adjacent faces for functions in $\mathbb{P}^q_{d-1}(\mathcal{F}_T)$. For any $T \in \mathcal{T}_H$, we denote by $(\Phi^{q,i}_T)_{1 \leqslant i \leqslant \mathrm{N}^q_d}$ a set of basis functions of the space $\mathbb{P}^q_d(T)$, and for any $F \in \mathcal{F}_H$, we denote by $(\Phi^{q,j}_F)_{1 \leqslant j \leqslant \mathrm{N}^q_{d-1}}$ a set of basis functions of the space $\mathbb{P}^q_{d-1}(F)$. We define, for any $T \in \mathcal{T}_H$ and $F \in \mathcal{F}_H$, Π^q_T and Π^q_F as the L^2-orthogonal projectors onto $\mathbb{P}^q_d(T)$ and $\mathbb{P}^q_{d-1}(F)$, respectively.

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

$$\|v\|_{L^2(F)} \leqslant c_{\mathrm{tr,d}} H_F^{-1/2} \|v\|_{L^2(T)} \qquad \forall v \in \mathbb{P}^q_d(T), \tag{7}$$

$$\|v\|_{L^2(F)} \leqslant c_{\mathrm{tr,c}} \left(H_T^{-1} \|v\|^2_{L^2(T)} + H_T \|\boldsymbol{\nabla} v\|^2_{L^2(T)^d} \right)^{1/2} \qquad \forall v \in H^1(T), \tag{8}$$

hold [15, Lemmas 1.46 and 1.49], as well as the local Poincaré inequality

$$\|v\|_{L^2(T)} \leqslant c_{\mathrm{P}} H_T \|\boldsymbol{\nabla} v\|_{L^2(T)^d} \quad \forall v \in H^1(T) \text{ such that } \int_T v = 0, \tag{9}$$

where $c_{\mathrm{P}} = \pi^{-1}$ for convex elements [7]; estimates in the nonconvex case can be found, e.g., in [39]. Finally, proceeding as in [24, Lemma 5.6], one can prove using the above trace and Poincaré inequalities that

$$|v - \Pi^q_T(v)|_{H^m(T)} + H_T^{1/2} |v - \Pi^q_T(v)|_{H^m(F)} \leqslant c_{\mathrm{app}} H_T^{s-m} |v|_{H^s(T)} \quad \forall v \in H^s(T), \tag{10}$$

for integers $1 \leqslant s \leqslant q+1$ and $0 \leqslant m \leqslant s$ (for $m = s$, (10) is a stability property). All of the above constants are independent of any meshsize and can depend on q, d, and on 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, the homogenized solution u_0, and the parameters α, β characterizing the spectrum of the diffusion matrices; the value of c can depend on the space dimension d, the underlying polynomial degree, the mesh regularity parameter γ, and some higher-order norms of the diffusion matrix $\mathbb{A}$ or the correctors μ_l that will be made clear from the context. A recent *hp*-analysis of the mono-scale HHO method can be found in [2].

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 multi-scale 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 $k \geqslant 1$. For all $1 \leqslant i \leqslant \mathrm{N}_d^{k-1}$, we consider the problem

$$\inf\left\{\int_T \left[\frac{1}{2}\mathbb{A}_\varepsilon \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\varphi - \Phi_T^{k-1,i}\,\varphi\right],\ \varphi \in H^1(T),\ \Pi_F^k(\varphi) = 0\ \forall F \in \mathcal{F}_T\right\}. \tag{11}$$

Problem (11) 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 $(\lambda_{F,j}^T)_{F\in\mathcal{F}_T,\,1\leqslant j\leqslant \mathrm{N}_{d-1}^k}$ 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 continuous problem

$$\left\{\begin{aligned} -\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla}\varphi_{\varepsilon,T}^{k+1,i}) &= \Phi_T^{k-1,i} && \text{in } T,\\ \mathbb{A}_\varepsilon \boldsymbol{\nabla}\varphi_{\varepsilon,T}^{k+1,i}\cdot\boldsymbol{n}_{T,F} &= \sum_{j=1}^{\mathrm{N}_{d-1}^k} \lambda_{F,j}^T \Phi_F^{k,j} && \text{on all } F \in \mathcal{F}_T,\\ \Pi_F^k(\varphi_{\varepsilon,T}^{k+1,i}) &= 0 && \text{for all } F \in \mathcal{F}_T. \end{aligned}\right. \tag{12}$$

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 $\varphi_{\varepsilon,T}^{k+1,i}$ for all $1 \leqslant i \leqslant \mathrm{N}_d^{k-1}$, 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 (12) approximately by using a classical (mono-scale) HHO method (or any other mono-scale approximation method). One can either consider a weak formulation in $\left\{\varphi \in H^1(T),\, \Pi_F^k(\varphi) = 0\, \forall F \in \mathcal{F}_T\right\}$, 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 (12) and (14) 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 $1 \leqslant j \leqslant \mathrm{N}_{d-1}^k$, we consider the problem

$$\inf \left\{ \int_T \left[\frac{1}{2} \mathbb{A}_\varepsilon \boldsymbol{\nabla} \varphi \cdot \boldsymbol{\nabla} \varphi \right],\, \varphi \in H^1(T),\, \Pi_F^k(\varphi) = \Phi_F^{k,j},\, \Pi_\sigma^k(\varphi) = 0\ \forall \sigma \in \mathcal{F}_T \backslash \{F\} \right\}. \tag{13}$$

Problem (13) 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 $(\lambda_{\sigma,q}^{T,F})_{\sigma \in \mathcal{F}_T,\, 1 \leqslant q \leqslant \mathrm{N}_{d-1}^k}$ 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 continuous problem

$$\begin{cases} -\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} \varphi_{\varepsilon,T,F}^{k+1,j}) = 0 & \text{in } T, \\ \mathbb{A}_\varepsilon \boldsymbol{\nabla} \varphi_{\varepsilon,T,F}^{k+1,j} \cdot \boldsymbol{n}_{T,\sigma} = \displaystyle\sum_{q=1}^{\mathrm{N}_{d-1}^k} \lambda_{\sigma,q}^{T,F} \Phi_\sigma^{k,q} & \text{on all } \sigma \in \mathcal{F}_T, \\ \Pi_F^k(\varphi_{\varepsilon,T,F}^{k+1,j}) = \Phi_F^{k,j}, & \\ \Pi_\sigma^k(\varphi_{\varepsilon,T,F}^{k+1,j}) = 0 & \text{for all } \sigma \in \mathcal{F}_T \backslash \{F\}. \end{cases} \tag{14}$$

4.2 Discrete space

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

$$V_{\varepsilon,T}^{k+1} := \left\{ v_\varepsilon \in H^1(T) \mid \mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon) \in \mathbb{P}_d^{k-1}(T),\, \mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}_{d-1}^k(\mathcal{F}_T) \right\}, \tag{15}$$

with the convention that $\mathbb{P}_d^{-1}(T) := \{0\}$. We recall that the condition $\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}_{d-1}^k(\mathcal{F}_T)$ is equivalent to $\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \cdot \boldsymbol{n}_{T,F} \in \mathbb{P}_{d-1}^k(F)$ for all $F \in \mathcal{F}_T$.

Proposition 4.2 (Characterization of $V_{\varepsilon,T}^{k+1}$). *For any $T \in \mathcal{T}_H$, the following holds:*

$$V_{\varepsilon,T}^{k+1} = \mathrm{Span}\left\{ (\varphi_{\varepsilon,T}^{k+1,i})_{1\leqslant i\leqslant \mathrm{N}_d^{k-1}}, (\varphi_{\varepsilon,T,F}^{k+1,j})_{F\in\mathcal{F}_T,\,1\leqslant j\leqslant \mathrm{N}_{d-1}^k} \right\}. \tag{16}$$

Moreover, the dimension of $V_{\varepsilon,T}^{k+1}$ is $\left(\mathrm{N}_d^{k-1} + \mathrm{card}(\mathcal{F}_T) \times \mathrm{N}_{d-1}^k\right)$ (or $\mathrm{card}(\mathcal{F}_T)$ if $k = 0$).

Proof. To establish (16), we only need to prove that

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

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

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

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 $\left(x_\sigma^{k,j}(v_\varepsilon)\right)_{1\leqslant j\leqslant \mathrm{N}_{d-1}^k}$ solve the linear system

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

It can be easily checked that $-\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla}\zeta) = 0$ in T and that $\mathbb{A}_\varepsilon \boldsymbol{\nabla}\zeta \cdot \boldsymbol{n}_{T,F} \in \mathbb{P}_{d-1}^k(F)$ and $\Pi_F^k(\zeta) = 0$ on all $F \in \mathcal{F}_T$. Using the compatibility conditions, we also infer that $\int_{\partial T} \mathbb{A}_\varepsilon \boldsymbol{\nabla}\zeta \cdot \boldsymbol{n}_{\partial T} = 0$, which means that the previous system for ζ is compatible. Hence, $\zeta \equiv 0$, which proves the converse inclusion. Finally, that the oscillatory basis functions are linearly independent can be shown by reasoning as above. $\square$

Remark 4.3 (Space $V^{k+1}_{\varepsilon,T}$)**.** *The definition of the space $V^{k+1}_{\varepsilon,T}$ is reminiscent of that considered in the non-conforming VE method in the case where $\mathbb{A}_\varepsilon = \mathbb{I}_d$; see [6] and also [13].*

We define $H_{\partial T} \in \mathbb{P}^0_{d-1}(\mathcal{F}_T)$ 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 $\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon$ for a function $v_\varepsilon \in V^{k+1}_{\varepsilon,T}$; for completeness, we also establish a bound on the divergence.

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

$$H_T \|\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon)\|_{L^2(T)} + \left\| H^{1/2}_{\partial T} \mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \cdot \boldsymbol{n}_{\partial T} \right\|_{L^2(\partial T)} \leqslant c\, \beta^{1/2} \left\| \mathbb{A}^{1/2}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \right\|_{L^2(T)^d}, \tag{17}$$

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 $d_\varepsilon := \mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon) \in \mathbb{P}^{k-1}_d(T)$. Let S be a simplicial sub-cell of T. Considering the standard bubble function $b_S \in H^1_0(S)$ (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,

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

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 $\|\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon)\|_{L^2(T)} \leqslant c\, \beta^{1/2} H_T^{-1} \left\| \mathbb{A}^{1/2}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \right\|_{L^2(T)^d}$. 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, $r_S := [\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon)]_{|S} \in \mathbb{P}^{k-1}_d(S) \subset \mathbb{P}^k_d(S)$ and $r_\sigma := [\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \cdot \boldsymbol{n}_{\partial T}]_{|\sigma} \in \mathbb{P}^k_{d-1}(\sigma)$. Note that $\boldsymbol{n}_{\partial T|\sigma} = \boldsymbol{n}_{\partial S|\sigma}$. Invoking [25, Lemma A.3], 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

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

with c' depending on γ (but not on k) and $\boldsymbol{H}(\mathrm{div};S) := \{\boldsymbol{z} \in L^2(S)^d \mid \mathrm{div}(\boldsymbol{z}) \in L^2(S)\}$. Since the function $[\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon]_{|S}$ 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 $\|\boldsymbol{q}\|_{L^2(S)^d} \leqslant c' \|\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon\|_{L^2(S)^d}$. A discrete trace inequality in the RTN finite element space shows that

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

where c depends on γ and k. We conclude by invoking mesh regularity. $\square$

4.3 Approximation properties

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

$$\begin{cases} -\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} \pi^{k+1}_{\varepsilon,T}(u_0)) = -\mathrm{div}(\mathbb{A}_0 \boldsymbol{\nabla} \Pi^{k+1}_T(u_0)) \in \mathbb{P}^{k-1}_d(T) & \text{in } T, \\ \mathbb{A}_\varepsilon \boldsymbol{\nabla} \pi^{k+1}_{\varepsilon,T}(u_0) \cdot \boldsymbol{n}_{\partial T} = \mathbb{A}_0 \boldsymbol{\nabla} \Pi^{k+1}_T(u_0) \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}^k_{d-1}(\mathcal{F}_T) & \text{on } \partial T. \end{cases} \tag{18}$$

Note that the data in (18) are compatible. From (18) we infer that, for any $w \in H^1(T)$,

$$\int_T \mathbb{A}_\varepsilon \boldsymbol{\nabla} \pi^{k+1}_{\varepsilon,T}(u_0) \cdot \boldsymbol{\nabla} w = \int_T \mathbb{A}_0 \boldsymbol{\nabla} \Pi^{k+1}_T(u_0) \cdot \boldsymbol{\nabla} w. \tag{19}$$

Lemma 4.5 (Approximation in $V^{k+1}_{\varepsilon,T}$)**.** *Assume that there is $\kappa > 0$ so that $\mathbb{A} \in C^{0,\kappa}(\mathbb{R}^d; \mathbb{R}^{d\times d})$ and that $u_0 \in H^{\max(k+2,3)}(T)$. Then,*

$$\begin{aligned} \|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla}(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0))\|_{L^2(T)^d} \leqslant c\, \beta^{1/2} \rho^{1/2} \Big(& H_T^{k+1} |u_0|_{H^{k+2}(T)} \\ & + \left(\varepsilon + (\varepsilon H_T)^{1/2}\right) |u_0|_{H^2(T)} + \varepsilon H_T |u_0|_{H^3(T)} + \varepsilon^{1/2} H_T^{-1/2} |u_0|_{H^1(T)} \Big), \end{aligned} \tag{20}$$

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

Proof. Subtracting/adding $\mathbb{A}_0 \boldsymbol{\nabla} u_0$ and using (19) with $w = \mathcal{L}^1_\varepsilon(u_0)_{|T} - \pi^{k+1}_{\varepsilon,T}(u_0)$ which is in $H^1(T)$, we infer that

$$\begin{aligned} \|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla}(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0))\|^2_{L^2(T)^d} = & \int_T (\mathbb{A}_\varepsilon \boldsymbol{\nabla} \mathcal{L}^1_\varepsilon(u_0) - \mathbb{A}_0 \boldsymbol{\nabla} u_0) \cdot \boldsymbol{\nabla}(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0)) \\ & + \int_T \mathbb{A}_0 \boldsymbol{\nabla}(u_0 - \Pi^{k+1}_T(u_0)) \cdot \boldsymbol{\nabla}(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0)). \end{aligned}$$

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

$$\begin{aligned} \|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla}(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0))\|_{L^2(T)^d} \leqslant & \beta^{1/2} \rho^{1/2} \|\boldsymbol{\nabla}(u_0 - \Pi^{k+1}_T(u_0))\|_{L^2(T)^d} \\ & + \alpha^{-1/2} \sup_{w \in H^1_\star(T)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\boldsymbol{\nabla} w\|_{L^2(T)^d}}, \end{aligned}$$

with $\mathcal{F}_\varepsilon(w) = \int_T (\mathbb{A}_\varepsilon \boldsymbol{\nabla} \mathcal{L}^1_\varepsilon(u_0) - \mathbb{A}_0 \boldsymbol{\nabla} u_0) \cdot \boldsymbol{\nabla} w$ and $H^1_\star(T) = \{w \in H^1(T) \mid \int_T w = 0\}$. The first term in the right-hand side is bounded using the approximation properties (10) of Π^{k+1}_T with $m = 1$ and $s = k+2$, and the second term is bounded in Lemma A.3. □

5 The MsHHO method

In this section, we introduce and analyze the multi-scale 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-based unknowns. In Section 5.3 we detail the solution strategy. 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 $k \geqslant 1$. For all $T \in \mathcal{T}_H$, we consider the following local set of discrete unknowns:

$$\underline{\mathrm{U}}_T^k := \mathbb{P}_d^{k-1}(T) \times \mathbb{P}_{d-1}^k(\mathcal{F}_T). \tag{21}$$

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 do not consider the case $k = 0$ since this corresponds to the method already analyzed in [33] (up to a slightly different treatment of the right-hand side; cf. Remark 5.4). 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 [13, 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 multi-scale 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 the well-posed local Neumann problem

$$\int_T \mathbb{A}_\varepsilon \boldsymbol{\nabla} p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) \cdot \boldsymbol{\nabla} w_\varepsilon = -\int_T \mathrm{v}_T \operatorname{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} w_\varepsilon) + \int_{\partial T} \mathrm{v}_{\mathcal{F}_T} \, \mathbb{A}_\varepsilon \boldsymbol{\nabla} w_\varepsilon \cdot \boldsymbol{n}_{\partial T} \quad \forall w_\varepsilon \in V_{\varepsilon,T}^{k+1}. \tag{22}$$

Note that (22) can be equivalently rewritten

$$\int_T \mathbb{A}_\varepsilon \boldsymbol{\nabla} p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) \cdot \boldsymbol{\nabla} w_\varepsilon = \int_T \boldsymbol{\nabla} \mathrm{v}_T \cdot \mathbb{A}_\varepsilon \boldsymbol{\nabla} w_\varepsilon - \int_{\partial T} (\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) \mathbb{A}_\varepsilon \boldsymbol{\nabla} w_\varepsilon \cdot \boldsymbol{n}_{\partial T} \quad \forall w_\varepsilon \in V_{\varepsilon,T}^{k+1}. \tag{23}$$

Integrating by parts the left-hand side of (22) and exploiting the definition (15) of the space $V_{\varepsilon,T}^{k+1}$, one can see that, for any $\underline{\mathrm{v}}_T \in \underline{\mathrm{U}}_T^k$,

$$\Pi_T^{k-1}\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right) = \Pi_T^{k-1}(\mathrm{v}_T) = \mathrm{v}_T, \qquad \Pi_{\partial T}^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right) = \Pi_{\partial T}^k(\mathrm{v}_{\mathcal{F}_T}) = \mathrm{v}_{\mathcal{F}_T}. \tag{24}$$

Owing to (15) and (22), we infer that, for all $v \in H^1(T)$,

$$\int_T \mathbb{A}_\varepsilon \boldsymbol{\nabla}\left(v - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k v)\right) \cdot \boldsymbol{\nabla} w_\varepsilon = 0 \quad \forall w_\varepsilon \in V_{\varepsilon,T}^{k+1}, \tag{25}$$

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

$$\left\|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla}\left(v - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k v)\right)\right\|_{L^2(T)^d} = \inf_{w_\varepsilon \in V_{\varepsilon,T}^{k+1}} \left\|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla}\left(v - w_\varepsilon\right)\right\|_{L^2(T)^d}. \tag{26}$$

Since the operator $p_{\varepsilon,T}^{k+1} \circ \underline{\mathrm{I}}_T^k$ preserves the mean value, its restriction to $V_{\varepsilon,T}^{k+1}$ is the identity operator.

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

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

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

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

$$\|\underline{\mathrm{v}}_T\|_T^2 := \|\boldsymbol{\nabla} \mathrm{v}_T\|_{L^2(T)^d}^2 + \left\|H_{\partial T}^{-1/2}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\right\|_{L^2(\partial T)}^2. \tag{27}$$

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

$$a_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T) \geqslant c\,\alpha \|\underline{\mathrm{v}}_T\|_T^2 \qquad \forall \underline{\mathrm{v}}_T \in \underline{\mathrm{U}}_T^k, \tag{28}$$

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

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

$$\begin{cases} -\mathrm{div}(\mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon) = -\triangle \mathrm{v}_T \in \mathbb{P}_d^{k-1}(T) & \text{in } T, \\ \mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \cdot \boldsymbol{n}_{\partial T} = \boldsymbol{\nabla} \mathrm{v}_T \cdot \boldsymbol{n}_{\partial T} \in \mathbb{P}_{d-1}^k(\mathcal{F}_T) & \text{on } \partial T, \end{cases} \tag{29}$$

and satisfying, e.g., $\int_T v_\varepsilon = 0$ (the way the constant is fixed is unimportant here). Note that data in (29) are compatible. Then, the following holds:

$$\int_T \mathbb{A}_\varepsilon \boldsymbol{\nabla} v_\varepsilon \cdot \boldsymbol{\nabla} z = \int_T \boldsymbol{\nabla} \mathrm{v}_T \cdot \boldsymbol{\nabla} z \qquad \forall z \in H^1(T).$$

Using this last relation where we take $z = p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)$, and using (23) where we take $w_\varepsilon = v_\varepsilon \in V^{k+1}_{\varepsilon,T}$ defined in (29), we infer that

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

After an integration by parts, this yields

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

By the Cauchy–Schwarz inequality and the discrete trace inequality (7), we then obtain

$$\|\boldsymbol{\nabla} \mathrm{v}_T\|_{L^2(T)^d} \leqslant c \left(\alpha^{-1/2} \|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla} p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d} + \left\| H_{\partial T}^{-1/2} (\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) \right\|_{L^2(\partial T)} \right). \tag{30}$$

To bound the second term in the right-hand side, we use (24) to infer that

$$\begin{aligned} [\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}]_{|\partial T} &= [\Pi^{k-1}_T \left(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) \right)]_{|\partial T} - \Pi^k_{\partial T} \left(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) \right) \\ &= \Pi^k_{\partial T} \left(\Pi^{k-1}_T \left(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) \right) - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) \right). \end{aligned}$$

Using the L^2-stability of $\Pi^k_{\partial T}$, the continuous trace inequality (8), the local Poincaré inequality (9) (since $p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) - \Pi^{k-1}_T \left(p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T) \right)$ has zero mean-value on T), and the H^1-stability of Π^{k-1}_T, we infer that

$$\left\| H_{\partial T}^{-1/2} (\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) \right\|_{L^2(\partial T)} \leqslant c\, \alpha^{-1/2} \|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla} p^{k+1}_{\varepsilon,T}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d}. \tag{31}$$

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

$$\mathbb{P}^{k-1}_d(\mathcal{T}_H) := \left\{ v \in L^2(\Omega) \mid v_{|T} \in \mathbb{P}^{k-1}_d(T)\ \forall T \in \mathcal{T}_H \right\}, \tag{32}$$

$$\mathbb{P}^k_{d-1}(\mathcal{F}_H) := \left\{ v \in L^2(\partial \mathcal{T}_H) \mid v_{|F} \in \mathbb{P}^k_{d-1}(F)\ \forall F \in \mathcal{F}_H \right\}. \tag{33}$$

The global set of discrete unknowns is defined to be

$$\underline{\mathrm{U}}^k_H := \mathbb{P}^{k-1}_d(\mathcal{T}_H) \times \mathbb{P}^k_{d-1}(\mathcal{F}_H), \tag{34}$$

so that any $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}^k_H$ can be decomposed as $\underline{\mathrm{v}}_H := (\mathrm{v}_{\mathcal{T}_H}, \mathrm{v}_{\mathcal{F}_H})$. For any given $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}^k_H$, we denote $\underline{\mathrm{v}}_T := (\mathrm{v}_T, \mathrm{v}_{\mathcal{F}_T}) \in \underline{\mathrm{U}}^k_T$ 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 $\underline{\mathrm{U}}_{H,0}^k := \{\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k \mid \mathrm{v}_F \equiv 0 \; \forall F \in \mathcal{F}_H^{\mathrm{b}}\}$. 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

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

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

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \int_\Omega f \mathrm{v}_{\mathcal{T}_H} \qquad \forall \underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k. \tag{35}$$

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 (27), we define 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:*

$$a_{\varepsilon,H}(\underline{\mathrm{v}}_H, \underline{\mathrm{v}}_H) = \sum_{T \in \mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla} p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right\|_{L^2(T)^d}^2 =: \|\underline{\mathrm{v}}_H\|_{\varepsilon,H}^2 \geqslant c\,\alpha \|\underline{\mathrm{v}}_H\|_H^2 \qquad \forall \underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k, \tag{36}$$

with constant c independent of ε, H, α and β. As a consequence, the discrete problem (35) 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*

$$V_{\varepsilon,H,0}^{k+1} := \left\{v_{\varepsilon,H} \in L^2(\Omega) \mid v_{\varepsilon,H|T} \in V_{\varepsilon,T}^{k+1} \; \forall\, T \in \mathcal{T}_H \; \textit{and} \; \Pi_F^k([\![v_{\varepsilon,H}]\!]_F) = 0 \; \forall\, F \in \mathcal{F}_H \right\},$$

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

$$\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) = \sum_{T \in \mathcal{T}_H} \int_T f\, \Pi_T^{k-1}(v_{\varepsilon,H}) \qquad \forall v_{\varepsilon,H} \in V_{\varepsilon,H,0}^{k+1}, \tag{37}$$

where $\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) := \sum_{T \in \mathcal{T}_H} \int_T \mathbb{A}_\varepsilon \boldsymbol{\nabla} u_{\varepsilon,H} \cdot \boldsymbol{\nabla} v_{\varepsilon,H}$. 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 (35) 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 (37). This proves that (35) is indeed a high-order (and polytopal) extension of the method in [33], up to a slightly different treatment of the right-hand side ($\Pi_T^{k-1}(v_{\varepsilon,H})$ instead of $v_{\varepsilon,H}$).

Let u_ε be the oscillatory solution to (1) and let $\underline{\mathrm{u}}_{\varepsilon,H}$ be the discrete MsHHO solution to (35). Let us define the discrete error such that

$$\underline{\mathrm{e}}_{\varepsilon,H} \in \underline{\mathrm{U}}_{H,0}^k, \qquad \underline{\mathrm{e}}_{\varepsilon,T} := \underline{\mathrm{I}}_T^k u_\varepsilon - \underline{\mathrm{u}}_{\varepsilon,T} \quad \forall T \in \mathcal{T}_H. \tag{38}$$

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 (38). Assume that $u_0 \in H^{k+2}(\Omega)$. Then, the following holds:*

$$\|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H} \leqslant c\,\rho^{1/2}\left(\beta \sum_{T\in\mathcal{T}_H} H_T^{2(k+1)} |u_0|^2_{H^{k+2}(T)} + \sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - \pi_{\varepsilon,T}^{k+1}(u_0)\right)\right\|^2_{L^2(T)^d}\right)^{1/2}, \tag{39}$$

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

Proof. Lemma 5.3 implies that

$$\|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H} = \sup_{\underline{\mathrm{v}}_H\in\underline{\mathrm{U}}_{H,0}^k} \frac{a_{\varepsilon,H}(\underline{\mathrm{e}}_{\varepsilon,H},\underline{\mathrm{v}}_H)}{\|\underline{\mathrm{v}}_H\|_{\varepsilon,H}}. \tag{40}$$

Let $\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k$. Performing an integration by parts, and using the facts that the flux $\mathbb{A}_0\boldsymbol{\nabla}u_0{\cdot}\boldsymbol{n}_F$ is continuous accross 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

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_{\varepsilon,H},\underline{\mathrm{v}}_H) = \int_\Omega f \mathrm{v}_{\mathcal{T}_H} = \sum_{T\in\mathcal{T}_H}\int_T \mathbb{A}_0\boldsymbol{\nabla}u_0{\cdot}\boldsymbol{\nabla}\mathrm{v}_T - \sum_{T\in\mathcal{T}_H}\int_{\partial T}(\mathrm{v}_T-\mathrm{v}_{\mathcal{F}_T})\mathbb{A}_0\boldsymbol{\nabla}u_0{\cdot}\boldsymbol{n}_{\partial T}. \tag{41}$$

Using (23), we then infer that

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

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

$$\begin{aligned}\mathfrak{T}_1 &= \sum_{T\in\mathcal{T}_H}\int_T \mathbb{A}_0\boldsymbol{\nabla}\left(\Pi_T^{k+1}(u_0)-u_0\right)\cdot\boldsymbol{\nabla}\mathrm{v}_T \\ &\quad - \sum_{T\in\mathcal{T}_H}\int_{\partial T}\mathbb{A}_0\boldsymbol{\nabla}\left(\Pi_T^{k+1}(u_0)-u_0\right)\cdot\boldsymbol{n}_{\partial T}(\mathrm{v}_T-\mathrm{v}_{\mathcal{F}_T}), \\ \mathfrak{T}_2 &= \sum_{T\in\mathcal{T}_H}\int_T\left(\mathbb{A}_\varepsilon\boldsymbol{\nabla}p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \mathbb{A}_0\boldsymbol{\nabla}\Pi_T^{k+1}(u_0)\right)\cdot\boldsymbol{\nabla}\mathrm{v}_T \\ &\quad - \sum_{T\in\mathcal{T}_H}\int_{\partial T}\left(\mathbb{A}_\varepsilon\boldsymbol{\nabla}p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \mathbb{A}_0\boldsymbol{\nabla}\Pi_T^{k+1}(u_0)\right){\cdot}\boldsymbol{n}_{\partial T}(\mathrm{v}_T-\mathrm{v}_{\mathcal{F}_T}).\end{aligned}$$

The term $\mathfrak{T}_1$ is estimated using Cauchy–Schwarz inequality and the approximation properties (10) of the projector Π_T^{k+1} for $m = 1$ and $s = k + 2$, yielding

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

Considering now $\mathfrak{T}_2$, we use the definition (18) of $\pi_{\varepsilon,T}^{k+1}(u_0)$ and the relation (19) to infer that

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

The first term in 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 $\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \pi_{\varepsilon,T}^{k+1}(u_0)\right) \in V_{\varepsilon,T}^{k+1}$. This yields

$$\begin{aligned} |\mathfrak{T}_2| & \leqslant c\,\beta^{1/2} \left(\sum_{T \in \mathcal{T}_H} \left\| \mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla} \left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon) - \pi_{\varepsilon,T}^{k+1}(u_0) \right) \right\|^2_{L^2(T)^d} \right)^{1/2} \|\underline{\mathrm{v}}_H\|_H \\ & \leqslant c\,\beta^{1/2} \left(\sum_{T \in \mathcal{T}_H} \left\| \mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla} \left(u_\varepsilon - \pi_{\varepsilon,T}^{k+1}(u_0) \right) \right\|^2_{L^2(T)^d} \right)^{1/2} \|\underline{\mathrm{v}}_H\|_H, \end{aligned}$$

where the last bound follows from (26) since $\pi_{\varepsilon,T}^{k+1}(u_0) \in V_{\varepsilon,T}^{k+1}$. Since $\|\underline{\mathrm{v}}_H\|^2_{\varepsilon,H} \geqslant c\,\alpha \|\underline{\mathrm{v}}_H\|^2_H$ owing to Lemma 5.3, we obtain the expected bound. □

Theorem 5.6 (Energy-error estimate)**.** *Assume that there is $\kappa > 0$ so that $\mathbb{A} \in C^{0,\kappa}(\mathbb{R}^d; \mathbb{R}^{d\times d})$ and that $u_0 \in H^{k+2}(\Omega)$ (recall that $k \geqslant 1$). Then, the following holds:*

$$\begin{aligned} \left(\sum_{T \in \mathcal{T}_H} \left\| \mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla} \left(u_\varepsilon - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{u}}_{\varepsilon,T}) \right) \right\|^2_{L^2(T)^d} \right)^{1/2} & \leqslant c\,\beta^{1/2} \rho \Bigg(\sum_{T \in \mathcal{T}_H} H_T^{2(k+1)} |u_0|^2_{H^{k+2}(T)} \\ + \varepsilon |\partial\Omega| |u_0|^2_{W^{1,\infty}(\Omega)} + \sum_{T \in \mathcal{T}_H} \Big[\left(\varepsilon^2 + \varepsilon H_T\right) |u_0|^2_{H^2(T)} & + \varepsilon^2 H_T^2 |u_0|^2_{H^3(T)} + \varepsilon H_T^{-1} |u_0|^2_{H^1(T)} \Big] \Bigg)^{1/2}, \end{aligned} \tag{42}$$

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 $\varepsilon \leqslant H \leqslant \ell_\Omega$ (ℓ_Ω denotes the diameter of Ω), we obtain an energy-error upper bound of the form $(\varepsilon^{1/2} + H^{k+1} + \varepsilon^{1/2} H^{-1/2})$.

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

$$\left(\sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla} e_{\varepsilon,T}\right\|^2_{L^2(T)^d}\right)^{1/2} \leqslant \left(\sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - p^{k+1}_{\varepsilon,T}(\underline{\mathrm{I}}^k_T u_\varepsilon)\right)\right\|^2_{L^2(T)^d}\right)^{1/2} + \left\|\underline{\mathrm{e}}_{\varepsilon,H}\right\|_{\varepsilon,H},$$

and owing to (26), we infer that

$$\left(\sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla} e_{\varepsilon,T}\right\|^2_{L^2(T)^d}\right)^{1/2} \leqslant \left(\sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - \pi^{k+1}_{\varepsilon,T}(u_0)\right)\right\|^2_{L^2(T)^d}\right)^{1/2} + \left\|\underline{\mathrm{e}}_{\varepsilon,H}\right\|_{\varepsilon,H}.$$

Lemma 5.5 then implies that

$$\left(\sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla} e_{\varepsilon,T}\right\|^2_{L^2(T)^d}\right)^{1/2} \leqslant \\ c\,\rho^{1/2}\left(\beta \sum_{T\in\mathcal{T}_H} H_T^{2(k+1)}|u_0|^2_{H^{k+2}(T)} + \sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - \pi^{k+1}_{\varepsilon,T}(u_0)\right)\right\|^2_{L^2(T)^d}\right)^{1/2}.$$

To conclude, we add/subtract $\mathcal{L}^1_\varepsilon(u_0)$ in the last term in the right-hand side, and invoke the triangle inequality together with Lemma A.4 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 $T \in \mathcal{T}_H$. □

Remark 5.7 (Dependency on ρ)**.** *The estimate (42) has a linear dependency with respect to the (global) heterogeneity/anisotropy ratio ρ (a close inspection of the proof shows that the term $\varepsilon^{1/2}|\partial\Omega|^{1/2}|u_0|_{W^{1,\infty}(\Omega)}$ only scales with $\rho^{1/2}$). This linear scaling is also obtained with the mono-scale HHO method when the diffusivity is non-constant in each mesh cell; cf. [18, Theorem 3.1].*

Remark 5.8 (Alternative estimate)**.** *It is possible to derive a different energy-error estimate under the slightly weaker regularity assumption that, for any $1 \leqslant l \leqslant d$, the corrector μ_l is in $W^{1,\infty}(\mathbb{R}^d)$. The assumption $u_0 \in H^{k+2}(\Omega)$ remains unchanged. As in [37], we then employ Lemma A.2 (with $D = T$) instead of Lemma A.3 in the proof of Lemma 4.5 yielding*

$$\left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}(\mathcal{L}^1_\varepsilon(u_0) - \pi^{k+1}_{\varepsilon,T}(u_0))\right\|_{L^2(T)^d} \leqslant \\ c\,\beta^{1/2}\rho^{1/2}\left(H_T^{k+1}|u_0|_{H^{k+2}(T)} + \varepsilon^{1/2}|\partial T|^{1/2}|u_0|_{W^{1,\infty}(T)} + \varepsilon|u_0|_{H^2(T)}\right).$$

The rest of the analysis is led as above, leading to the following energy-error estimate in *lieu of* (42)*:*

$$\left(\sum_{T\in\mathcal{T}_H}\left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{u}}_{\varepsilon,T})\right)\right\|_{L^2(T)^d}^2\right)^{1/2} \leqslant c\,\beta^{1/2}\rho\Bigg(\sum_{T\in\mathcal{T}_H} H_T^{2(k+1)}|u_0|_{H^{k+2}(T)}^2$$
$$+\,\varepsilon|\partial\Omega||u_0|_{W^{1,\infty}(\Omega)}^2 + \varepsilon^2|u_0|_{H^2(\Omega)}^2\Bigg)^{1/2},$$

which essentially leads to a behavior of the form $(\varepsilon^{1/2}+H^{k+1})$ *for* $\varepsilon\leqslant H\leqslant \ell_\Omega$*. This upper bound is less sharp than that derived in Theorem 5.6 in the sense that it does not capture the resonance phenomenon observed numerically when the meshsize is not too large with respect to* ε*.*

5.2 The equal-order case

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

$$\underline{\mathrm{U}}_T^k := \mathbb{P}_d^k(T)\times\mathbb{P}_{d-1}^k(\mathcal{F}_T). \tag{43}$$

Any element $\underline{\mathrm{v}}_T\in\underline{\mathrm{U}}_T^k$ 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 [13, Section 2.4], it can be proved that, for all $T\in\mathcal{T}_H$, the restriction of $\underline{\mathrm{I}}_T^k$ to $\tilde V_{\varepsilon,T}^{k+1}$ is an isomorphism from $\tilde V_{\varepsilon,T}^{k+1}$ to $\underline{\mathrm{U}}_T^k$, where

$$\tilde V_{\varepsilon,T}^{k+1} := \left\{v_\varepsilon\in H^1(T)\mid \operatorname{div}(\mathbb{A}_\varepsilon\boldsymbol{\nabla}v_\varepsilon)\in\mathbb{P}_d^k(T),\ \mathbb{A}_\varepsilon\boldsymbol{\nabla}v_\varepsilon\cdot\boldsymbol{n}_{\partial T}\in\mathbb{P}_{d-1}^k(\mathcal{F}_T)\right\}. \tag{44}$$

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

The local multi-scale reconstruction operator $p_{\varepsilon,T}^{k+1} : \underline{\mathrm{U}}_T^k\to V_{\varepsilon,T}^{k+1}$ is still defined as in (22), so that the key relations (25) and (26) still hold. In particular, $p_{\varepsilon,T}^{k+1}\circ\underline{\mathrm{I}}_T^k : H^1(T)\to V_{\varepsilon,T}^{k+1}$ is the $\mathbb{A}_\varepsilon$-weighted elliptic projection. However, the restriction of $p_{\varepsilon,T}^{k+1}\circ\underline{\mathrm{I}}_T^k$ to the larger space $\tilde V_{\varepsilon,T}^{k+1}$ is *not* the identity operator since $p_{\varepsilon,T}^{k+1}$ maps onto the smaller space $V_{\varepsilon,T}^{k+1}$. Concerning (24), we still have $\Pi_{\partial T}^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right) = \mathrm{v}_{\mathcal{F}_T}$, but now $\Pi_T^{k-1}\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right) = \Pi_T^{k-1}(\mathrm{v}_T)$ is in general different from v_T. This leads us to introduce the symmetric, positive semi-definite stabilization

$$j_{\varepsilon,T}(\underline{\mathrm{u}}_T,\underline{\mathrm{v}}_T) := \alpha\int_{\partial T} H_{\partial T}^{-1}\left(\mathrm{u}_T - \Pi_T^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{u}}_T)\right)\right)\left(\mathrm{v}_T - \Pi_T^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right)\right). \tag{45}$$

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

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

Remark 5.9 (Variant). *Alternatively, one can discard the stabilization at the prize of computing additional oscillatory cell-based basis functions, using the basis functions* $(\Phi_T^{k,i})_{1\leqslant i\leqslant \mathrm{N}_d^k}$ *instead of* $(\Phi_T^{k-1,i})_{1\leqslant i\leqslant \mathrm{N}_d^{k-1}}$ *as proposed in Section 4.1.1. This is the approach pursued in [34] (up to a slightly different treatment of the right-hand side) for* $k = 0$ *where one oscillatory cell-based basis function is added (in the slightly different context of perforated domains). The analysis for polynomial degrees* $k \geqslant 1$ *is similar to the one presented in Section 5.1 and is omitted for brevity.*

Recall the local stability semi-norm $\|\cdot\|_T$ defined by (27).

Lemma 5.10 (Local stability and approximation). *The following holds:*

$$a_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T) \geqslant c\,\alpha\|\underline{\mathrm{v}}_T\|_T^2 \qquad \forall \underline{\mathrm{v}}_T \in \underline{\mathrm{U}}_T^k. \tag{46}$$

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

$$j_{\varepsilon,T}(\underline{\mathrm{I}}_T^k v, \underline{\mathrm{I}}_T^k v)^{1/2} \leqslant c\left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(v - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k v)\right)\right\|_{L^2(T)^d}, \tag{47}$$

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 (30) on $\|\boldsymbol{\nabla}\mathrm{v}_T\|_{L^2(T)^d}$ still holds, so that we only need to bound $\left\|H_{\partial T}^{-1/2}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\right\|_{L^2(\partial T)}$. Since $\Pi_{\partial T}^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right) = \mathrm{v}_{\mathcal{F}_T}$, we infer that $(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T}) = \Pi_{\partial T}^k\left(\mathrm{v}_T - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right)$, so that invoking the L^2-stability of $\Pi_{\partial T}^k$ and the triangle inequality while adding/subtracting $\Pi_T^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right)$, we obtain

$$\begin{aligned}\left\|H_{\partial T}^{-1/2}(\mathrm{v}_T - \mathrm{v}_{\mathcal{F}_T})\right\|_{L^2(\partial T)} &\leqslant \left\|H_{\partial T}^{-1/2}\left(\mathrm{v}_T - \Pi_T^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right)\right)\right\|_{L^2(\partial T)} \\ &\quad + \left\|H_{\partial T}^{-1/2}\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T) - \Pi_T^k\left(p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right)\right)\right\|_{L^2(\partial T)}.\end{aligned}$$

The first term in the right-hand side is bounded by $\alpha^{-1/2}j_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T)^{1/2}$, 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 (31)) by $c\,\alpha^{-1/2}\left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\right\|_{L^2(T)^d}$. To prove (47), we start from

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

The result then follows from the application of the discrete trace inequality (7), of the L^2-stability property of Π_T^k, and of the local Poincaré inequality (9) (since $\int_T p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k v) = \int_T v$). □

We define the broken polynomial space

$$\mathbb{P}_d^k(\mathcal{T}_H) := \left\{v \in L^2(\Omega) \mid v_{|T} \in \mathbb{P}_d^k(T)\ \forall T \in \mathcal{T}_H\right\},$$

and the global set of discrete unknowns is defined to be

$$\underline{\mathrm{U}}_H^k := \mathbb{P}_d^k(\mathcal{T}_H) \times \mathbb{P}_{d-1}^k(\mathcal{F}_H), \tag{48}$$

where $\mathbb{P}_{d-1}^k(\mathcal{F}_H)$ is still defined by (33). To take into account homogeneous Dirichlet boundary conditions, we consider again the subspace $\underline{\mathrm{U}}_{H,0}^k := \{\underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k \mid \mathrm{v}_F \equiv 0 \; \forall F \in \mathcal{F}_H^{\mathrm{b}}\}$. 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

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

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

$$a_{\varepsilon,H}(\underline{\mathrm{u}}_{\varepsilon,H}, \underline{\mathrm{v}}_H) = \int_\Omega f \mathrm{v}_{\mathcal{T}_H} \qquad \forall \underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_{H,0}^k. \tag{49}$$

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:*

$$\begin{aligned} a_{\varepsilon,H}(\underline{\mathrm{v}}_H, \underline{\mathrm{v}}_H) &= \sum_{T \in \mathcal{T}_H} \left(\|\mathbb{A}_\varepsilon^{1/2} \boldsymbol{\nabla} p_{\varepsilon,T}^{k+1}(\underline{\mathrm{v}}_T)\|_{L^2(T)^d}^2 + j_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T) \right) \\ &=: \|\underline{\mathrm{v}}_H\|_{\varepsilon,H}^2 \geqslant c\,\alpha \|\underline{\mathrm{v}}_H\|_H^2 \qquad \forall \underline{\mathrm{v}}_H \in \underline{\mathrm{U}}_H^k, \end{aligned} \tag{50}$$

with constant c independent of ε, H, α and β. As a consequence, the discrete problem (49) is well-posed.

Remark 5.12 (NcFE interpretation)**.** *As in Remark 5.4, it is possible to give a NcFE interpretation of the scheme (49). Let*

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

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

$$\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) = \sum_{T \in \mathcal{T}_H} \int_T f\, \Pi_T^k(v_{\varepsilon,H}) \qquad \forall v_{\varepsilon,H} \in \tilde{V}_{\varepsilon,H,0}^{k+1}, \tag{51}$$

where $\tilde{a}_{\varepsilon,H}(u_{\varepsilon,H}, v_{\varepsilon,H}) := \sum_{T \in \mathcal{T}_H} a_{\varepsilon,T}\left(\underline{\mathrm{I}}_T^k(u_{\varepsilon,H|T}), \underline{\mathrm{I}}_T^k(v_{\varepsilon,H|T})\right)$. Then, it can be shown that $\underline{\mathrm{u}}_{\varepsilon,H}$ solves (49) 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 (51). 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 $p_{\varepsilon,T}^{k+1} \circ \underline{\mathrm{I}}_T^k$ to $\tilde{V}_{\varepsilon,T}^{k+1}$ is not the identity operator. As in the mono-scale HHO method, the operator $p_{\varepsilon,T}^{k+1}$, which maps onto the smaller space $V_{\varepsilon,T}^{k+1}$, allows one to restrict the number of computed basis functions while maintaining optimal (and here also ε-robust) approximation properties. The basis functions (from the discrete space $\tilde{V}_{\varepsilon,T}^{k+1}$) 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 (38). Assume that $u_0 \in H^{k+2}(\Omega)$. Then, the following holds:*

$$\|\underline{\mathrm{e}}_{\varepsilon,H}\|_{\varepsilon,H} \leqslant c\,\rho^{1/2}\left(\beta \sum_{T\in\mathcal{T}_H} H_T^{2(k+1)}|u_0|^2_{H^{k+2}(T)} + \sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - \pi_{\varepsilon,T}^{k+1}(u_0)\right)\right\|^2_{L^2(T)^d}\right)^{1/2}, \quad (52)$$

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

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

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

$$\begin{aligned}|\mathfrak{T}_3| &\leqslant \left(\sum_{T\in\mathcal{T}_H} j_{\varepsilon,T}\left(\underline{\mathrm{I}}_T^k u_\varepsilon, \underline{\mathrm{I}}_T^k u_\varepsilon\right)\right)^{1/2}\left(\sum_{T\in\mathcal{T}_H} j_{\varepsilon,T}(\underline{\mathrm{v}}_T, \underline{\mathrm{v}}_T)\right)^{1/2} \\ &\leqslant c\left(\sum_{T\in\mathcal{T}_H} \left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{I}}_T^k u_\varepsilon)\right)\right\|^2_{L^2(T)^d}\right)^{1/2} \|\underline{\mathrm{v}}_H\|_{\varepsilon,H},\end{aligned}$$

where we have used (47). We can now conclude as before. □

Theorem 5.14 (Energy-error estimate). *Assume that there is $\kappa > 0$ so that $\mathbb{A} \in C^{0,\kappa}(\mathbb{R}^d;\mathbb{R}^{d\times d})$ and that $u_0 \in H^{\max(k+2,3)}(\Omega)$. Then, the following holds:*

$$\left(\sum_{T\in\mathcal{T}_H}\left\|\mathbb{A}_\varepsilon^{1/2}\boldsymbol{\nabla}\left(u_\varepsilon - p_{\varepsilon,T}^{k+1}(\underline{\mathrm{u}}_{\varepsilon,T})\right)\right\|^2_{L^2(T)^d}\right)^{1/2} \leqslant c\,\beta^{1/2}\rho\Bigg(\sum_{T\in\mathcal{T}_H} H_T^{2(k+1)}|u_0|^2_{H^{k+2}(T)}$$
$$+\,\varepsilon|\partial\Omega||u_0|^2_{W^{1,\infty}(\Omega)} + \sum_{T\in\mathcal{T}_H}\left[\left(\varepsilon^2 + \varepsilon H_T\right)|u_0|^2_{H^2(T)} + \varepsilon^2 H_T^2|u_0|^2_{H^3(T)} + \varepsilon H_T^{-1}|u_0|^2_{H^1(T)}\right]\Bigg)^{1/2}, \quad (53)$$

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 $\varepsilon \leqslant H \leqslant \ell_\Omega$, we obtain an energy-error upper bound of the form $(\varepsilon^{1/2} + H^{k+1} + \varepsilon^{1/2}H^{-1/2})$.

Proof. Identical to that of Theorem 5.6. □

Remark 5.15 (Dependency on ρ). *As in the mixed-order case (cf. Remark 5.7), the estimate (53) has a linear dependency with respect to the (global) heterogeneity/anisotropy ratio ρ.*

Remark 5.16 (Alternative estimate). *An alternative estimate to (53) can be derived in the spirit of Remark 5.8.*

5.3 Offline/online solution strategy

Let us consider the equal-order version ($k \geqslant 0$) of the MsHHO method introduced in Section 5.2. Similar considerations carry over to the mixed-order case ($k \geqslant 1$) of Section 5.1. To solve (49), 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 (12) (cf. Remark 4.1), and the $\mathrm{card}(\mathcal{F}_T) \times \mathrm{N}_{d-1}^k$ functions $\varphi_{\varepsilon,T,F}^{k+1,j}$ solution to (14). This first substep is fully parallelizable. In a second time, we compute the multi-scale reconstruction operators $p_{\varepsilon,T}^{k+1}$, by solving (22) for all $T \in \mathcal{T}_H$. Each computation requires to invert a symmetric positive-definite matrix of size $\left(\mathrm{N}_d^{k-1} + \mathrm{card}(\mathcal{F}_T) \times \mathrm{N}_{d-1}^k\right)$, 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 [18, Section 3.3.1]. Basically, in each cell, this substep consists in inverting a symmetric positive-definite matrix of size N_d^k (N_d^{k-1} when solving (35)). 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 $\mathbb{P}_d^k(\mathcal{T}_H)$ ($\mathbb{P}_d^{k-1}(\mathcal{T}_H)$ when solving (35)), 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) for another f (or for other boundary conditions), only the online step must be rerun.

For the implementation of the mono-scale HHO method, we refer to [11].

A Estimates on the first-order two-scale expansion

In this appendix, we derive various useful estimates on the first-order two-scale expansion $\mathcal{L}_\varepsilon^1(u_0)$ defined by (6). Except for Lemma A.3, these estimates are classical; we provide (short) proofs since we additionally track the 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

$$w \mapsto \mathcal{F}_\varepsilon(w) := \int_D \left(\mathbb{A}_\varepsilon \boldsymbol{\nabla} \mathcal{L}_\varepsilon^1(u_0) - \mathbb{A}_0 \boldsymbol{\nabla} u_0\right) \cdot \boldsymbol{\nabla} w, \tag{54}$$

for all $w \in H^1_0(D)$ (Dirichlet case), or for all $w \in H^1_\star(D) := \{w \in H^1(D) \mid \int_D w = 0\}$ (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 $1 \leqslant l \leqslant d$, the corrector μ_l belongs to $W^{1,\infty}(\mathbb{R}^d)$. Then,*

$$\sup_{w \in H^1_0(D)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\boldsymbol{\nabla} w\|_{L^2(D)^d}} \leqslant c\,\beta\varepsilon |u_0|_{H^2(D)}, \tag{55}$$

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

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

$$\begin{aligned}
\left[\mathbb{A}_\varepsilon \boldsymbol{\nabla} \mathcal{L}^1_\varepsilon(u_0)\right]_i &= \sum_{j=1}^d [\mathbb{A}_\varepsilon]_{ij} \partial_j \mathcal{L}^1_\varepsilon(u_0) \\
&= \sum_{j=1}^d [\mathbb{A}_\varepsilon]_{ij} \left(\partial_j u_0 + \varepsilon \sum_{l=1}^d \left(\frac{1}{\varepsilon} \mathcal{R}_\varepsilon(\partial_j \mu_l) \partial_l u_0 + \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0 \right) \right) \\
&= [\mathbb{A}_0 \boldsymbol{\nabla} u_0]_i + \sum_{l=1}^d \mathcal{R}_\varepsilon(\theta^l_i) \partial_l u_0 + \varepsilon \sum_{l,j=1}^d [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l) \partial^2_{j,l} u_0,
\end{aligned} \tag{56}$$

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:

- $\theta^l_i \in L^\infty_{\rm per}(Q)$ by assumption on $\mathbb{A}$ and on the correctors μ_l;
- $\int_Q \theta^l_i = 0$ as a consequence of (3);
- $\sum_{i=1}^d \partial_i \theta^l_i = 0$ in $\mathbb{R}^d$ as a consequence of (4).

Adapting [31, Equation (1.11)] (see also [27, Sections I.3.1 and I.3.3]), we infer that, for any integer $1 \leqslant l \leqslant d$, there exists a skew-symmetric matrix $\mathbb{T}^l \in W^{1,\infty}_{\rm per}(Q)^{d \times d}$, satisfying $\int_Q \mathbb{T}^l = \mathbb{0}$ and such that, for any integer $1 \leqslant i \leqslant d$,

$$\theta^l_i = \sum_{q=1}^d \partial_q \mathbb{T}^l_{qi}. \tag{57}$$

Plugging (57) into (56), we infer that, for any integer $1 \leqslant i \leqslant d$,

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

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 (54) of $\mathcal{F}_\varepsilon$, this yields

$$\mathcal{F}_\varepsilon(w) = \varepsilon\Bigg(\sum_{i,l,j=1}^d \int_D [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l)\partial^2_{j,l} u_0\, \partial_i w - \sum_{i,l,q=1}^d \int_D \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial^2_{q,l} u_0\, \partial_i w \Bigg) + \varepsilon \sum_{i,l,q=1}^d \int_D \partial_q \left(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0\right) \partial_i w. \tag{58}$$

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

$$\mathcal{F}_\varepsilon(w) = \varepsilon\Bigg(\sum_{i,l,j=1}^d \int_D [\mathbb{A}_\varepsilon]_{ij} \mathcal{R}_\varepsilon(\mu_l)\partial^2_{j,l} u_0\, \partial_i w - \sum_{i,l,q=1}^d \int_D \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial^2_{q,l} u_0\, \partial_i w \Bigg) + \varepsilon \sum_{i,l,q=1}^d \int_{\partial D} \partial_q \left(\mathcal{R}_\varepsilon(\mathbb{T}^l_{qi})\partial_l u_0\right) n_{\partial D,i}\, w, \tag{59}$$

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

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

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

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

$$\sup_{w \in H^1_\star(D)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\boldsymbol{\nabla} w\|_{L^2(D)^d}} \leqslant c\,\beta \left(\varepsilon |u_0|_{H^2(D)} + |\partial D|^{1/2} \varepsilon^{1/2} |u_0|_{W^{1,\infty}(D)} \right), \tag{60}$$

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

Proof. Our starting point is (58). The first two terms in 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 [31, p. 29], we define, for $\eta > 0$, the domain $D_\eta := \{\boldsymbol{x} \in D \mid \mathrm{dist}(\boldsymbol{x}, \partial D) < \eta\}$. 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 $\zeta_\eta(\boldsymbol{x}) = \mathrm{dist}(\boldsymbol{x}, \partial D)/\eta$ if $\boldsymbol{x} \in D_\eta$, and $\zeta_\eta(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in D \backslash D_\eta$. We have $0 \leqslant \zeta_\eta \leqslant 1$ and $\max_{1 \leqslant q \leqslant d} \|\partial_q \zeta_\eta\|_{L^\infty(D)} \leqslant \eta^{-1}$. We first infer that

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

since $(1-\zeta_\eta)$ vanishes identically on $D \backslash D_\eta$ and since $\sum_{i,l,q=1}^{d} \int_D \partial_q \left(\zeta_\eta \mathcal{R}_\varepsilon(\mathbb{T}^l_{qi}) \partial_l u_0\right) \partial_i w = 0$ as can be seen by integration by parts, using the fact that $\mathbb{T}^l_{qi} = -\mathbb{T}^l_{iq}$ for any integers $1 \leqslant i, q \leqslant d$, and the fact that ζ_η vanishes identically on ∂D. Then, accounting for the fact that

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

we infer that

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

Using the estimate $|D_\eta| \leqslant \eta |\partial D|$, the fact that $\max_{1 \leqslant l \leqslant d} \|\mathbb{T}^l\|_{W^{1,\infty}(\mathbb{R}^d)^{d \times d}} \leqslant c\,\beta$, and since the function $\eta \mapsto \frac{\varepsilon}{\sqrt{\eta}} + \sqrt{\eta}$ is minimal for $\eta = \varepsilon$, we finally infer the bound (60). ☐

Lemma A.3 (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,*

$$\sup_{w \in H^1_\star(D)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\boldsymbol{\nabla} w\|_{L^2(D)^d}} \leqslant c\,\beta \left(\left(\varepsilon + (\varepsilon \ell_D)^{1/2}\right) |u_0|_{H^2(D)} + \varepsilon \ell_D |u_0|_{H^3(D)} + \varepsilon^{1/2} \ell_D^{-1/2} |u_0|_{H^1(D)} \right), \tag{61}$$

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

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 $\iota > 0$ as the Hölder continuity of $\mathbb{A}$ on $\mathbb{R}^d$ implies the Hölder continuity of μ_l and $\boldsymbol{\nabla} \mu_l$ on $\mathbb{R}^d$ for any $1 \leqslant l \leqslant d$; cf., e.g., [26, Theorem 8.22 and Corollary 8.36]. Following [31, p. 6-7] and [33, p. 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 (59). The first two terms in 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

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

Using the Cauchy–Schwarz inequality and the trace inequality (8), the first term in the right-hand side can be estimated as

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

since $\max_{1\leqslant l\leqslant d}\left\|\mathbb{T}^l\right\|_{C^0(\mathbb{R}^d)^{d\times d}} \leqslant c\,\beta$. Observing that $\int_D w = 0$, we can use the Poincaré inequality (9) to infer that

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

To estimate the second term in the right-hand side, we adapt the ideas from [33, 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

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

where the vectors $\boldsymbol{\tau}^{qi}_\sigma$ are such that $\|\boldsymbol{\tau}^{qi}_\sigma\|_{\ell^2} \leqslant 1$ and $\boldsymbol{\tau}^{qi}_\sigma\cdot\boldsymbol{n}_{\partial D|\sigma} = 0$. Then, using a straightforward adaptation of the result in [33, Lemma 4.6], and since $\max_{1\leqslant l\leqslant d}\left\|\mathbb{T}^l\right\|_{C^1(\mathbb{R}^d)^{d\times d}} \leqslant c\,\beta$, we infer that

$$\left|\int_\sigma \mathcal{R}_\varepsilon\left(\boldsymbol{\nabla}\mathbb{T}^l_{qi}\right)\cdot\boldsymbol{\tau}^{qi}_\sigma\,\partial_l u_0\, w\right| \leqslant c\,\beta\varepsilon^{1/2}H_S^{-3/2}\left(|u_0|_{H^1(S)} + H_S|u_0|_{H^2(S)}\right) \\ \left(\|w\|_{L^2(S)} + H_S\|\boldsymbol{\nabla} w\|_{L^2(S)^d}\right),$$

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

$$|\mathfrak{T}_2| \leqslant c\,\beta\varepsilon^{1/2}\ell_D^{-3/2}\left(|u_0|_{H^1(D)} + \ell_D|u_0|_{H^2(D)}\right)\left(\|w\|_{L^2(D)} + \ell_D\|\boldsymbol{\nabla} w\|_{L^2(D)^d}\right).$$

Finally, invoking the Poincaré inequality (9) since w has zero mean-value in D yields

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

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

A.2 Global energy-norm estimate

Lemma A.4 (Energy-norm estimate)**.** *Assume that the homogenized solution u_0 belongs to $W^{1,\infty}(\Omega) \cap H^2(\Omega)$, and that, for any $1 \leqslant l \leqslant d$, the corrector μ_l belongs to $W^{1,\infty}(\mathbb{R}^d)$. Then,*

$$\left\| \mathbb{A}_\varepsilon^{1/2} \nabla (u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0)) \right\|_{L^2(\Omega)^d} \leqslant c\, \beta^{1/2} \left(|\partial\Omega|^{1/2} \varepsilon^{1/2} \, |u_0|_{W^{1,\infty}(\Omega)} + \rho^{1/2} \varepsilon \, |u_0|_{H^2(\Omega)} \right), \tag{62}$$

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

Proof. The regularity assumptions on u_0 and the correctors imply $(u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0)) \in H^1(\Omega)$; however, we do not have $(u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0)) \in H_0^1(\Omega)$. Following the ideas in [31, p. 28], we define, for $\eta > 0$, the domain $\Omega_\eta := \{ \boldsymbol{x} \in \Omega \mid \mathrm{dist}(\boldsymbol{x}, \partial\Omega) < \eta \}$. 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 $\zeta_\eta(\boldsymbol{x}) = \mathrm{dist}(\boldsymbol{x}, \partial\Omega)/\eta$ if $\boldsymbol{x} \in \Omega_\eta$, and $\zeta_\eta(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in \Omega \backslash \Omega_\eta$. We have $0 \leqslant \zeta_\eta \leqslant 1$ and $\max\limits_{1 \leqslant i \leqslant d} \|\partial_i \zeta_\eta\|_{L^\infty(\Omega)} \leqslant \eta^{-1}$. 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:

$$\begin{aligned} \left\| \mathbb{A}_\varepsilon^{1/2} \nabla (u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0)) \right\|_{L^2(\Omega)^d} \leqslant{} & \left\| \mathbb{A}_\varepsilon^{1/2} \nabla (u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0)) \right\|_{L^2(\Omega)^d} \\ & + \left\| \mathbb{A}_\varepsilon^{1/2} \nabla (\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0)) \right\|_{L^2(\Omega)^d}. \end{aligned} \tag{63}$$

Let us focus on the first term in the right-hand side of (63). We have

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

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

$$\begin{aligned} \left\| \mathbb{A}_\varepsilon^{1/2} \nabla (u_\varepsilon - \mathcal{L}_\varepsilon^{1,0}(u_0)) \right\|_{L^2(\Omega)^d} \leqslant{} & \alpha^{-1/2} \sup_{w \in H_0^1(\Omega)} \frac{\left| \int_\Omega \mathbb{A}_\varepsilon \nabla \left(u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0) \right) \cdot \nabla w \right|}{\|\nabla w\|_{L^2(\Omega)^d}} \\ & + \left\| \mathbb{A}_\varepsilon^{1/2} \nabla (\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0)) \right\|_{L^2(\Omega)^d}. \end{aligned} \tag{64}$$

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) and (5), the estimates (63) and (64) lead to

$$\begin{aligned} \left\| \mathbb{A}_\varepsilon^{1/2} \nabla (u_\varepsilon - \mathcal{L}_\varepsilon^1(u_0)) \right\|_{L^2(\Omega)^d} \leqslant{} & \alpha^{-1/2} \sup_{w \in H_0^1(\Omega)} \frac{|\mathcal{F}_\varepsilon(w)|}{\|\nabla w\|_{L^2(\Omega)^d}} \\ & + 2\beta^{1/2} \left\| \nabla (\mathcal{L}_\varepsilon^1(u_0) - \mathcal{L}_\varepsilon^{1,0}(u_0)) \right\|_{L^2(\Omega)^d}, \end{aligned} \tag{65}$$

recalling that $\mathcal{F}_\varepsilon(w) = \int_\Omega (\mathbb{A}_\varepsilon \nabla \mathcal{L}^1_\varepsilon(u_0) - \mathbb{A}_0 \nabla u_0) \cdot \nabla w$. Since we can bound the first term in the right-hand side of (65) using Lemma A.1 (with $D = \Omega$), it remains to estimate the second term. Owing to the definition of ζ_η, we infer that

$$\left\| \nabla (\mathcal{L}^1_\varepsilon(u_0) - \mathcal{L}^{1,0}_\varepsilon(u_0)) \right\|_{L^2(\Omega)^d} = \varepsilon \left\| \nabla \left((1-\zeta_\eta) \sum_{l=1}^d \mathcal{R}_\varepsilon(\mu_l) \partial_l u_0 \right) \right\|_{L^2(\Omega_\eta)^d}. \tag{66}$$

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

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

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

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

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

Remark A.5 (Weaker regularity assumption)**.** *Without the regularity assumption $u_0 \in W^{1,\infty}(\Omega)$, one can still invoke a Sobolev embedding since $u_0 \in H^2(\Omega)$. The first term between the parentheses in the right-hand side of* (62) *becomes $c(\Omega,p)|\partial\Omega|^{1/2-1/p}\varepsilon^{1/2-1/p}(\ell_\Omega^{-1}|u_0|_{H^1(\Omega)} + |u_0|_{H^2(\Omega)})$ where $p = 6$ for $d = 3$ and p can be taken as large as wanted for $d = 2$ (note that $c(\Omega,p) \to +\infty$ when $p \to +\infty$ in that case). We refer, e.g., to [37] for the derivation of estimates in this setting.*

References

[1] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The Heterogeneous Multiscale Method. *Acta Numerica*, 21:1–87, 2012.

[2] J. Aghili, D. A. Di Pietro, and B. Ruffini. A *hp*-Hybrid High-Order method for variable diffusion on general meshes. Submitted, 2016. Available at hal-01290251.

[3] G. Allaire. *Shape Optimization by the Homogenization Method*, volume 146 of *Applied Mathematical Sciences*. Springer, New York, 2002.

[4] R. Araya, C. Harder, D. Paredes, and F. Valentin. Multiscale Hybrid-Mixed method. *SIAM J. Numer. Anal.*, 51(6):3505–3531, 2013.

[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(5):1749–1779, 2002.

[6] B. Ayuso de Dios, K. Lipnikov, and G. Manzini. The nonconforming virtual element method. *ESAIM: Math. Model Numer. Anal. (M2AN)*, 50(3):879–904, 2016.

[7] M. Bebendorf. A note on the Poincaré inequality for convex domains. *Z. Anal. Anwendungen*, 22(4):751–756, 2003.

[8] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *Math. Models Methods Appl. Sci. (M3AS)*, 23:199–214, 2013.

[9] A. Cangiani, E. H. Georgoulis, and P. Houston. *hp*-version discontinuous Galerkin methods on polygonal and polyhedral meshes. *Math. Models Methods Appl. Sci. (M3AS)*, 24(10):2009–2041, 2014.

[10] E. T. Chung, S. Fu, and Y. Yang. An enriched multiscale mortar space for high contrast flow problems. Submitted, 2016. Available at arXiv:1609.02610.

[11] M. Cicuttin, D. A. Di Pietro, and A. Ern. Implementation of Discontinuous Skeletal methods on arbitrary-dimensional, polytopal meshes using generic programming. Submitted, 2017. Available at hal-01429292.

[12] B. Cockburn. Static condensation, hybridization, and the devising of the HDG methods. In G. R. Barrenechea, F. Brezzi, A. Cangiani, and E. H. Georgoulis, editors, *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, number 114 in Lecture Notes in Computational Science and Engineering, pages 129–177. Springer, 2016.

[13] 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(3):635–650, 2016.

[14] 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(2):1319–1365, 2009.

[15] D. A. Di Pietro and A. Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69 of *Mathématiques & Applications*. Springer-Verlag, Berlin, 2012.

[16] 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*, 353:31–34, 2015.

[17] 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(4):461–472, 2014.

[18] D. A. Di Pietro, A. Ern, and S. Lemaire. A review of Hybrid High-Order methods: formulations, computational aspects, comparison with other methods. In G. R. Barrenechea, F. Brezzi, A. Cangiani, and E. H. Georgoulis, editors, *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, volume 114 of *Lecture Notes in Computational Science and Engineering*, pages 205–236. Springer, 2016.

[19] W. E and B. Engquist. The Heterogeneous Multiscale Methods. *Comm. Math. Sci.*, 1:87–132, 2003.

[20] Y. Efendiev and T. Y. Hou. *Multiscale Finite Element Methods - Theory and Applications*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer-Verlag, New York, 2009.

[21] Y. Efendiev, T. Y. Hou, and X.-H. Wu. Convergence of a nonconforming multiscale finite element method. *SIAM J. Numer. Anal.*, 37(3):888–910, 2000.

[22] 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.*, 292:243–256, 2015.

[23] 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(1):342–369, 2015.

[24] A. Ern and J.-L. Guermond. Finite element quasi-interpolation and best approximation. *ESAIM: Math. Model Numer. Anal. (M2AN)*, 2017. DOI: 10.1051/m2an/2016066.

[25] 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. Submitted, 2016. Available at hal-01422204.

[26] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[27] V. Girault and P.-A. Raviart. *Finite element methods for Navier-Stokes equations*, volume 5 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, 1986. Theory and algorithms.

[28] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comp. Physics*, 134:169–189, 1997.

[29] 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(227):913–943, 1999.

[30] 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(2):185–205, 2004.

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

[32] R. Kornhuber and H. Yserentant. Numerical homogenization of elliptic multiscale problems by subspace decomposition. *Multiscale Model. Simul.*, 14(3):1017–1036, 2016.

[33] C. Le Bris, F. Legoll, and A. Lozinski. MsFEM à la Crouzeix–Raviart for highly oscillatory elliptic problems. *Chinese Annals of Mathematics, Series B*, 34(1):113–138, 2013.

[34] C. Le Bris, F. Legoll, and A. Lozinski. An MsFEM-type approach for perforated domains. *SIAM Multiscale Modeling and Simulation*, 12(3):1046–1077, 2014.

[35] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83:2583–2603, 2014.

[36] L. Mu, J. Wang, and X. Ye. A Weak Galerkin generalized multiscale finite element method. *J. Comp. Appl. Math.*, 305:68–81, 2016.

[37] D. Paredes, F. Valentin, and H. M. Versieux. On the robustness of Multiscale Hybrid-Mixed methods. *Math. Comp.*, 86:525–548, 2017.

[38] N. Sukumar and A. Tabarraei. Conforming polygonal finite elements. *Int. J. Numer. Methods Engrg.*, 61(12):2045–2066, 2004.

[39] A. Veeser and R. Verfürth. Poincaré constants for finite element stars. *IMA J. Numer. Anal.*, 32(1):30–47, 2012.

[40] E. L. Wachspress. *A Rational Finite Element Basis*, volume 114 of *Mathematics in Science and Engineering.* Academic Press, 1975.

[41] J. Wang and X. Ye. A weak Galerkin finite element method for second-order elliptic problems. *J. Comput. Appl. Math.*, 241:103–115, 2013.