ESAIM: M2AN 59 (2025) 1213–1237
https://doi.org/10.1051/m2an/2025029

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

A LOCALIZED ORTHOGONAL DECOMPOSITION STRATEGY FOR HYBRID DISCONTINUOUS GALERKIN METHODS

Peipei Lu[1], Roland Maier[2,*] and Andreas Rupp[3]

Abstract. We formulate and analyze a multiscale method for an elliptic problem with an oscillatory coefficient based on a skeletal (hybrid) formulation. More precisely, we employ hybrid discontinuous Galerkin approaches and combine them with the localized orthogonal decomposition methodology to obtain a coarse-scale skeletal method that effectively includes fine-scale information. This work is the first step in reliably merging hybrid skeletal formulations and localized orthogonal decomposition to unite the advantages of both strategies. Numerical experiments are presented to illustrate the theoretical findings.

Mathematics Subject Classification. 65N12, 65N30.

Received July 10, 2024. Accepted April 3, 2025.

1. Introduction

Multiscale problems are difficult because oscillation scales need to be resolved when using classical numerical methods such as finite element methods. Otherwise, suboptimal convergence rates and pre-asymptotic effects can be observed. However, globally, resolving fine-scale features is computationally unfeasible. Moreover, often, only the coarse-scale behavior of solutions is of interest. Computational multiscale methods tackle this problem and construct problem-adapted approximation spaces that suitably include fine-scale information while operating on a coarse scale of interest.

We present a hybrid multiscale method that approximately solves elliptic multiscale problems in the framework of the localized orthogonal decomposition method (LOD). We aim to combine the advantages of hybrid finite elements (local discretization character, rigorous analysis for static condensation, enhanced convergence properties) with the advantages of the LOD (rigorous multiscale and localization analysis as well as optimal rates in the multiscale setting under minimal structural assumptions). This work provides an initial analysis and is a first step towards enhanced hybrid, local, and high-order discretizations of multiscale problems based on the LOD.

Prominent (conforming) multiscale methods that only pose minimal structural assumptions on the diffusion coefficient in the context of elliptic multiscale problems are, *e.g.*, generalized (multiscale) finite element meth-

Keywords and phrases. Multiscale method, hybrid method, elliptic problems, Poincaré–Friedrichs inequalities for DG and HDG.

[1] Department of Mathematics Sciences, Soochow University, Suzhou 215006, P.R. China.

[2] Institute for Applied and Numerical Mathematics, Karlsruhe Institute of Technology, Englerstr. 2, 76131 Karlsruhe, Germany.

[3] Department of Mathematics, Saarland University, 66123 Saarbrücken, Germany.

*Corresponding author: roland.maier@kit.edu

ods [8–10, 24, 44] (which are particularly well-suited for high-contrast problems), adaptive local bases [32, 56], the LOD [36, 47], gamblets [52], and their variants. The minimal assumptions on the coefficient come at the moderate cost of an increased computational overhead (*e.g.*, increased support of the basis functions or an increased number of basis functions per mesh entity, typically logarithmically concerning the mesh size). For a more detailed overview of corresponding methods, we refer to the textbooks [48, 53] and the review article [4]. Recently, progress has been made regarding the question whether the computational overhead can be further reduced, see [33]. The vast majority of the (early) multiscale methods restrict themselves to lowest-order approaches, which appears to be a natural choice because of the typical lack of regularity of the solution. However, there are also approaches that obtain higher-order convergence properties. Examples are the methods in [2, 39] based on the heterogeneous multiscale method or in [3, 37] based on the multiscale finite element method. These methods, however, require additional structural assumptions on the coefficient.

Generally, nonconforming finite elements, such as the discontinuous Galerkin (DG) method, seem to be well-suited for higher-order approximations as they relax the continuity constraint in their test and trial spaces and are, *e.g.*, employed for the multiscale strategies in [1, 55]. The relaxed continuity constraint also gives hope to retaining a very localized influence of small-scale structures. The first DG method in connection with the LOD was presented in [26, 27], where a truly DG multiscale space of first order is constructed, but an extension to a higher-order setting seems possible. Besides, even in the original LOD method, discontinuous functions are beneficial when localizing the computations, see [36]. A conforming LOD-type multiscale method that is based on DG spaces for suitable constraint conditions is investigated in [23, 46] (see also [45]) and achieves higher-order convergence rates under minimal regularity assumptions, only requiring piecewise regularity of the right-hand side. Regarding the super-localization with higher-order rates, we refer to [29]. An LOD approach in connection with mixed finite elements is presented in [35].

Compared to classical DG methods, hybrid finite element methods use additional unknowns on the skeleton of the mesh, which offers several desirable properties. They have been introduced in [28] and allow for static condensation, which reduces the number of unknowns in the linear system of equations and recovers a symmetric positive matrix for mixed methods. Later, Arnold and Brezzi [7] discovered that the additional information in the skeleton unknowns can be used to enhance the order of convergence in a postprocessing step. A comprehensive overview of several hybrid finite element methods such as hybrid Raviart–Thomas, hybrid Brezzi–Douglas–Marini, and other hybrid discontinuous Galerkin methods can be found in [19]. Another popular class of hybrid finite elements is the hybrid high-order (HHO) method initially proposed in [22]. Compared to the hybrid methods, HHO methods use a more involved stabilization that grants improved convergence properties.

The attractive properties of hybrid finite elements have led to several strategies to use them in a multiscale context, such as multiscale hybrid-mixed methods [5, 34] and the multiscale hybrid high-order method [17] (that can be shown to be equivalent, see [14]). Another approach is the multiscale hybrid method in [11] related to the multiscale hybrid-mixed method. These methods introduce oscillating shape functions in the spirit of a multiscale finite element approach, similar to the ideas in [38] based on Crouzeix–Raviart elements. Other hybrid multiscale approaches are, *e.g.*, the multiscale mortar mixed finite element method [6] or the multiscale HDG method [25]. These methods typically require additional (piecewise) regularity assumptions on the coefficient, at least concerning the coarse scale of interest. Finally, the hybrid localized spectral decomposition approach in [49] is another hybrid multiscale method, focusing on high-contrast problems.

Our method is the first hybrid finite element scheme based on the LOD methodology. It requires minimal assumptions on the coefficient and uses a fully condensed setting. We construct oscillating shape functions on the skeleton (union of mesh faces) and not in the bulk domain (union of mesh elements). LOD-based multiscale methods typically rely on the nestedness of discrete spaces on different mesh levels. Such a property does not hold for (fully) condensed hybrid methods since test and trial spaces are constructed on the skeleton of the mesh, which grows as the mesh is refined. This obstacle also appears if one considers multigrid methods to solve or precondition the systems of equations that arise from fully condensed hybrid finite element methods. At the moment, there are two competing approaches to tackle this issue: heterogeneous multigrid methods (see *e.g.*, [20]), whose first step is to replace the discretization scheme with a non-hybrid scheme, and homogeneous

multigrid methods (see *e.g.*, [40–43]), which use the same discretization on all mesh levels. Some of our analyses rely on techniques developed using the latter approach. Although both homogeneous and heterogeneous multigrid methods have certain advantages, in particular in the context of multiscale problems, we believe that homogeneous methods are better suited for LOD approaches since they exhibit more regular execution patterns and the numerical approximates have the same basic properties (mass conservation, etc.) on all levels.

A particular advantage of our mixed hybrid approach over primal finite element formulations is the in situ generation of an optimally convergent flux field. This flux field can be post-processed into an optimally converging $H(\mathrm{div})$-conforming flow field that is of primary interest for reactive transport for porous media if the considered elliptic problem (see (2.1) below) is interpreted as Darcy's equation. On the contrary, non-hybridized dual mixed methods induce a saddle-point structure that is usually harder to solve than our statically condensed system of equations.

The remaining parts of this paper are structured as follows. Section 2 presents the model problem and introduces HDG methods. We then provide beneficial preliminary results in Section 3.2 and present a prototypical skeletal multiscale method based on the ideas of the LOD in Section 4. The method is analyzed in Section 5, and a localized version is discussed in Section 6. Numerical illustrations are presented in Section 7. Some auxiliary results on a lifting operator and a Poincaré–Friedrichs inequality for broken skeletal spaces are presented in the appendix.

2. Problem formulation

2.1. Elliptic model problem

We investigate the diffusion equation in dual form. That is, we search for $u\colon \Omega \to \mathbb{R}$ and $\boldsymbol{q}\colon \Omega \to \mathbb{R}^d$ (in suitable spaces) such that

$$\begin{aligned} \operatorname{div} \boldsymbol{q} &= f && \text{in } \Omega, \\ A^{-1}\boldsymbol{q} + \nabla u &= 0 && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{2.1}$$

with respect to some bounded Lipschitz polytope $\Omega \subseteq \mathbb{R}^d$, $d \in \{1,2,3\}$. Here, $f \in L^2(\Omega)$ denotes a given right hand side and $A \in \mathfrak{A}$ is a rough coefficient with

$$\mathfrak{A} := \left\{ \begin{array}{l} A \in L^\infty\big(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}}\big) \colon \exists\, 0 < \alpha \le \beta < \infty : \\ \forall \xi \in \mathbb{R}^d \text{ and a.e. } x \in \Omega \colon \alpha|\xi|^2 \le A(x)\xi \cdot \xi \le \beta|\xi|^2 \end{array} \right\}. \tag{2.2}$$

In particular, A is symmetric, bounded, and uniformly elliptic. We are especially interested in the case in which A may describe ε-scale features of some given medium, where $0 < \varepsilon \ll 1$.

2.2. Meshes

The unknown u in (2.1) can usually not be computed analytically. Thus, we need to approximate it numerically. To this end, we use a hybrid discontinuous Galerkin (HDG) method. In this work, the numerical scheme is based on the degrees of freedom of a coarse mesh $\mathcal{E}_H$ but also relies on information on a fine mesh $\mathcal{E}_h$. The respective sets of faces are denoted by $\mathcal{F}_H$ and $\mathcal{F}_h$, and the union of all faces is denoted the *skeleton*. For the sake of simplicity, we assume both meshes to be geometrically conforming, shape regular, and simplicial, and that $\mathcal{E}_h$ can be generated from $\mathcal{E}_H$ by uniform refinement. In what follows, we use the placeholder $\star$ to indicate that the fine or the coarse mesh size could be considered, *i.e.*, $\star \in \{h, H\}$.

The fine mesh $\mathcal{E}_h$ is assumed to be fine enough to resolve the small-scale information encoded in the coefficient. For simplicity, we assume that A is an element-wise constant with respect to an intermediate mesh $\mathcal{E}_\varepsilon$, *i.e.*,

$$A|_e = c_e > 0 \qquad \text{for all } e \in \mathcal{E}_\varepsilon, \tag{2.3}$$

where $h < \varepsilon < H$, and we again assume that the corresponding meshes are refinements of each other. In particular, A is also element-wise constant on the mesh $\mathcal{E}_h$. In the present setting, ε is relatively small indicating that A contains heavy oscillations. Therefore, $\mathcal{E}_h$ possibly includes many elements such that a direct global finite element computation for it is unfeasible. That is why one would like to conduct numerical simulations concerning the coarser mesh $\mathcal{E}_H$. Note that throughout this work, we denote mesh elements of $\mathcal{E}_h$ by e and mesh elements of $\mathcal{E}_H$ by E. If both are possible, we use E.

3. HDG methods

In this manuscript, we consider several classes of HDG methods (LDG-H, EDG, RT-H, and BDM-H methods) of degree at most p and construct a localized orthogonal decomposition framework that can be used in concert with any of these HDG methods. To this end, we start with an abstract definition of HDG methods and derive properties (of all described HDG methods) that can be used for localized orthogonal decomposition in this section. Notably, although HDG methods are generally of order p, the convergence rate of the resulting multiscale method is only one due to the minimal assumptions regarding regularity.

3.1. Definition of the considered HDG finite element methods

Rendering a unified analysis, we use the approach of [19] and characterize the hybrid methods by defining their local test and trial spaces and their local solvers. For $\star \in \{h, H\}$, the test and trial spaces are locally given by

- M_F for the skeletal unknown $m_\star|_F$ approximating the trace of u on $F \in \mathcal{F}_\star$,
- V_E for the primal unknown $u_\star|_E$ approximating u in $E \in \mathcal{E}_\star$,
- $\boldsymbol{W}_E$ for the dual unknown $\boldsymbol{q}_\star|_E$ approximating $\boldsymbol{q}$ in $E \in \mathcal{E}_\star$,

and they can be concatenated to respective global spaces, *i.e.*,

$$
\begin{aligned}
M_\star &:= \left\{ m \in L^2(\mathcal{F}_\star) \;\middle|\; \begin{array}{ll} m|_F \in M_F & \text{for all } F \in \mathcal{F}_\star \\ m|_F = 0 & \text{for all } F \subset \partial\Omega \end{array} \right\}, \\
V_\star &:= \left\{ v \in L^2(\Omega) \,\middle|\, v|_E \in V_E \quad \text{for all } E \in \mathcal{E}_\star \right\}, \\
\boldsymbol{W}_\star &:= \left\{ \boldsymbol{q} \in L^2\big(\Omega; \mathbb{R}^d\big) \,\middle|\, \boldsymbol{q}|_E \in \boldsymbol{W}_E \quad \text{for all } E \in \mathcal{E}_\star \right\}.
\end{aligned}
$$

Note that here and in the following sections, we slightly abuse the notation and write $L^2(\mathcal{F}_\star)$, which needs to be understood face-wise. For later use, we also define the localized skeletal space on the interior of a union of elements $\omega \subset \Omega$ by

$$
M_\star(\omega) := \left\{ m \in L^2(\mathcal{F}_\star \cap \overline{\omega}) \;\middle|\; \begin{array}{ll} m|_F \in M_F & \text{for all } F \in \mathcal{F}_\star \cap \overline{\omega} \\ m|_F = 0 & \text{for all } F \subset \partial\omega \end{array} \right\}.
$$

If $\omega = \Omega$, we have $M_\star = M_\star(\Omega)$. This kind of definition applies verbatim to other localized spaces. Note that we may use that functions in $M_\star(\omega)$ can be extended by zero to the whole domain Ω.

The local solvers to determine a discrete solution on the whole domain are characterized by the mappings

$$
\begin{aligned}
\mathcal{U} &\colon L^2(\mathcal{F}_h) \to V_h, & \mathcal{V} &\colon L^2(\Omega) \to V_h, \\
\boldsymbol{\mathcal{Q}} &\colon L^2(\mathcal{F}_h) \to \boldsymbol{W}_h, & \boldsymbol{\mathcal{P}} &\colon L^2(\Omega) \to \boldsymbol{W}_h
\end{aligned}
$$

that will be defined in more detail below. We emphasize that our multiscale strategy will only require these local solvers on the fine scale (*i.e.*, with mesh parameter h). Therefore, we will not use a subscript h when we refer to the local solvers. Notably, the left-hand side operators are usually restricted to

$$
\mathcal{U} \colon M_h \to V_h \quad \text{and} \quad \boldsymbol{\mathcal{Q}} \colon M_h \to \boldsymbol{W}_h,
$$

and they are defined element-wise. For a given element $e \in \mathcal{E}_h$, they map the skeletal unknown m to u_e and $\boldsymbol{q}_e$ satisfying

$$\int_e A^{-1}\boldsymbol{q}_e \cdot \boldsymbol{p}_e \,\mathrm{d}x - \int_e u_e \operatorname{div} \boldsymbol{p}_e \,\mathrm{d}x = -\int_{\partial e} m\boldsymbol{p}_e \cdot \boldsymbol{\nu} \,\mathrm{d}\sigma \tag{3.1a}$$

$$\int_{\partial e} (\boldsymbol{q}_e \cdot \boldsymbol{\nu} + \tau u_e) v_e \,\mathrm{d}\sigma - \int_e \boldsymbol{q}_e \cdot \nabla v_e \,\mathrm{d}x = \tau \int_{\partial e} m v_e \,\mathrm{d}\sigma \tag{3.1b}$$

for all $v_e \in V_e$, $\boldsymbol{p}_e \in \boldsymbol{W}_e$ and some $\tau \geq 0$ (which will be set later). These local operators, and those local operators that will be defined later, are well-defined since the resulting linear systems of equations describing (3.1) are square, and $m = 0$ implies $u_e = 0$ and $\boldsymbol{q}_e = 0$ if the LDG-H, EDG, RT-H, or BDM-H methods are selected (see the end of this section for definitions of the respective methods). In this sense $\mathcal{U}\colon m \mapsto u_e$, $\boldsymbol{\mathcal{Q}} : m \mapsto \boldsymbol{q}_e$ for all $e \in \mathcal{E}_h$. Importantly, (3.1b) can be rewritten as

$$\int_{\partial e} \boldsymbol{q}_e \cdot \boldsymbol{\nu} v_e \,\mathrm{d}\sigma - \int_e \boldsymbol{q}_e \cdot \nabla v_e \,\mathrm{d}x = -\underbrace{\int_{\partial e} \tau(u_e - m) v_e \,\mathrm{d}\sigma}_{=s_e(m,v)} \tag{3.2}$$

and we have for $\mu \in L^2(\mathcal{F}_h)$ that

$$\mathcal{U}\mu = \mathcal{U}\pi_h\mu \qquad \text{and} \qquad \boldsymbol{\mathcal{Q}}\mu = \boldsymbol{\mathcal{Q}}\pi_h\mu, \tag{3.3}$$

where π_h is the face-wise L^2-orthogonal projection onto M_h.

The local solvers $\mathcal{V}$ and $\boldsymbol{\mathcal{P}}$ for the right-hand side are defined analogously but do not play a crucial role in our analysis. We refer to [19] for further details. If m is the HDG approximate of the trace of u, we have

$$\mathcal{U}m + \mathcal{V}f = u_h \approx u \qquad \text{and} \qquad \boldsymbol{\mathcal{Q}}m + \boldsymbol{\mathcal{P}}f = \boldsymbol{q}_h \approx \boldsymbol{q}.$$

With these abstract concepts at hand, we can characterize an HDG method as a method that seeks $m \in M_h$ such that

$$a(m, \mu) = \int_\Omega f\,\mathcal{U}\mu \,\mathrm{d}x \tag{3.4}$$

for all $\mu \in M_h$, where

$$a(m, \mu) = \int_\Omega A^{-1}\boldsymbol{\mathcal{Q}}m \cdot \boldsymbol{\mathcal{Q}}\mu \,\mathrm{d}x + \underbrace{\sum_{e\in\mathcal{E}_h} \tau \int_{\partial e} (\mathcal{U}m - m)(\mathcal{U}\mu - \mu) \,\mathrm{d}\sigma}_{=s(m,\mu)} \tag{3.5}$$

is the HDG bilinear form and s is called *penalty* or *stabilizer*. Again, we highlight that the bilinear form a lives on the fine mesh. Although it is, in principle, also possible to transfer all concepts to the coarse mesh, this is not required in our work.

The different types of HDG methods that are covered by our analysis can be distinguished by the following choices:

- For the *hybridized local discontinuous Galerkin* (**LDG-H**) method, we choose all local spaces to consist of polynomials of degree at most p, *i.e.*, $M_F = \mathbb{P}_p(F)$, $V_E = \mathbb{P}_p(E)$, and $\boldsymbol{W}_E = \mathbb{P}_p^d(E)$, where $E \in \mathcal{E}_\star$ and $F \in \mathcal{F}_\star$. In our work, the parameter $\tau > 0$ is chosen as a face-wise constant such that $\tau h \lesssim 1$. Common choices for such τ are $\tau = \frac{1}{h}$ and $\tau = 1$.
- The *embedded discontinuous Galerkin* (**EDG**) method uses the choices of the LDG-H method but additionally requires $M_\star$ to be overall continuous.
- By setting $\tau = 0$ in the LDG-H method and replacing the local bulk spaces with the RT-H and BDM-H spaces, we obtain the *hybridized Raviart–Thomas* (**RT-H**) and the *hybridized Brezzi–Douglas–Marini* (**BDM-H**) methods, respectively.

Note that we write $b \lesssim c$ if there exists a constant $C > 0$ (that might depend on the dimension) such that $b \leq C\,c$. However, we track the dependencies on the upper and lower bounds of A, α and β, explicitly.

3.2. Auxiliary definitions

We start by defining the space of localized, overall continuous, piece-wise affine-linear finite elements on $\omega \subset \Omega$,

$$\overline{V}_\star(\overline{\omega}) = \{v \in C(\omega)\colon v|_E \in \mathbb{P}_1(E) \ \text{ for all } E \in \mathcal{E}_\star \text{ with } E \subset \overline{\omega}, \text{ and } v|_{\partial\omega} = 0\}$$

and its trace space

$$\overline{M}_\star(\omega) = \gamma_\star \overline{V}_\star(\omega) = \{\mu \in C(\mathcal{F}_\star \cap \overline{\omega})\colon \mu|_F \in \mathbb{P}_1(F) \ \text{ for all } F \in \mathcal{F}_\star \cap \overline{\omega},\ \mu|_{\partial\omega} = 0\},$$

which is a subspace of $M_\star(\omega)$. Note that $\gamma_\star$ is the trace operator. As above, we abbreviate $\overline{V}_\star := \overline{V}_\star(\Omega)$ and $\overline{M}_\star := \overline{M}_\star(\Omega)$. Again, the localized spaces can be globalized by extending their entries by zero. We also emphasize that the space $\overline{M}_\star$ is exactly the first-order EDG test and trial space. Next, we define a mesh-dependent L^2-type scalar product (and a corresponding norm) suitable for the above skeletal spaces. These scalar products will be important for the error analysis later on. For functions $m, \mu \in L^2(\mathcal{F}_\star)$, we set

$$\langle m, \mu\rangle_\star = \sum_{E\in\mathcal{E}_\star} \frac{|E|}{|\partial E|} \int_{\partial E} m\mu \,\mathrm{d}\sigma \tag{3.6}$$

with $\star \in \{h, H\}$ as above. This scalar product readily extends to the case in which m and μ are replaced by bulk functions $u, v \in V_\star$ (where we consider the integral over the traces of u, v). If additionally $u, v \in \overline{V}_\star$, the definition is commensurate to the L^2 scalar product in the bulk. Further, we define the mesh-dependent norm $\|\cdot\|_\star = \sqrt{\langle\cdot,\cdot\rangle_\star}$ and denote with $\|\cdot\|_0$ the standard L^2 norm for bulk functions. Moreover, we define for any union of (either coarse or fine) elements $\omega \subset \Omega$, the local norm

$$\|m\|^2_{\star,\omega} = \sum_{\substack{E'\in\mathcal{E}_\star,\\ E'\subset\omega}} \frac{|E'|}{|\partial E'|} \sum_{\substack{F\in\mathcal{F}_\star,\\ F\cap E\neq\emptyset}} \int_{F\cap\partial E'} m^2 \,\mathrm{d}\sigma.$$

Similarly, $\|\cdot\|_{0,E}$ denotes the localized L^2 norm on E. For later use, we define the element patch ω_E for $E \in \mathcal{E}_\star$ as

$$\omega_E = \operatorname{int}\Big\{x \in \Omega \mid x \in E' \in \mathcal{E}_\star \text{ with } \overline{E}' \cap \overline{E} \neq \emptyset\Big\}. \tag{3.7}$$

We also define larger element patches iteratively by

$$\mathsf{N}^1(E) := \omega_E, \qquad \mathsf{N}^\ell(E) := \mathsf{N}^1\big(\mathsf{N}^{\ell-1}(E)\big), \qquad \ell \in \mathbb{N}, \ell \geq 2. \tag{3.8}$$

Further, we denote with $\nabla_\star v \in L^2(\Omega)$ the broken gradient of an element-wise defined function $v \in L^2(\Omega)$ with $v \in H^1(E)$ for all $E \in \mathcal{E}_\star$. More precisely, we have $(\nabla_\star u)|_E = \nabla(u|_E)$ for all $E \in \mathcal{E}_\star$.

3.3. Properties of HDG

Given the above construction of the local HDG solvers, we can introduce a useful identity for later calculations.

Lemma 3.1. *Let $\xi, \rho \in L^2(\mathcal{F}_h)$. Then*

$$\int_e A^{-1}\boldsymbol{\mathcal{Q}}\xi \cdot \boldsymbol{\mathcal{Q}}\rho \,\mathrm{d}x = -\int_{\partial e} (\boldsymbol{\mathcal{Q}}\rho \cdot \boldsymbol{\nu})\,\xi + \tau(\mathcal{U}\rho - \rho)\,\mathcal{U}\xi \,\mathrm{d}\sigma$$

for any $e \in \mathcal{E}_h$.

Proof. Given $e \in \mathcal{E}_h$, we have that

$$\begin{aligned}\int_e A^{-1}\boldsymbol{\mathcal{Q}}\xi \cdot \boldsymbol{\mathcal{Q}}\rho \,\mathrm{d}x &= \int_e \mathcal{U}\xi(\operatorname{div}\boldsymbol{\mathcal{Q}}\rho)\,\mathrm{d}x - \int_{\partial e} \xi\,(\boldsymbol{\mathcal{Q}}\rho\cdot\boldsymbol{\nu})\,\mathrm{d}\sigma\\ &= -\int_e \nabla\mathcal{U}\xi\cdot\boldsymbol{\mathcal{Q}}\rho\,\mathrm{d}x + \int_{\partial e}(\mathcal{U}\xi-\xi)\boldsymbol{\mathcal{Q}}\rho\cdot\boldsymbol{\nu}\,\mathrm{d}\sigma\\ &= \int_{\partial e}(\boldsymbol{\mathcal{Q}}\rho\cdot\boldsymbol{\nu})\mathcal{U}\xi\,\mathrm{d}\sigma - \int_e \boldsymbol{\mathcal{Q}}\rho\cdot\nabla\mathcal{U}\xi\,\mathrm{d}x - \int_{\partial e}(\boldsymbol{\mathcal{Q}}\rho\cdot\boldsymbol{\nu})\xi\,\mathrm{d}\sigma\\ &= -s_e(\rho,\mathcal{U}\xi) - \int_{\partial e}(\boldsymbol{\mathcal{Q}}\rho\cdot\boldsymbol{\nu})\,\xi\,\mathrm{d}\sigma,\end{aligned}$$

where the first equality is (3.1a) with $\boldsymbol{p} = \boldsymbol{\mathcal{Q}}\rho$ and the second one is integration by parts. The last equation combines the two terms containing $\mathcal{U}\xi$ to the penalty term by exploiting (3.2) with $v_e = \mathcal{U}\xi$ and $\boldsymbol{q}_e = \boldsymbol{\mathcal{Q}}\rho$. □

The local solvers of all HDG methods of Section 3 also fulfill the following essential conditions for any $\mu \in M_h$ and $e \in \mathcal{E}_h$. These properties are generalizations of their analogs shown in [41] and will be heavily used in the following sections.

– The trace of the bulk unknown approximates the skeletal unknown, *i.e.*,

$$\|\mathcal{U}\mu - \mu\|_h \lesssim h\big\|A^{-1}\boldsymbol{\mathcal{Q}}\mu\big\|_0 \leq \tfrac{h}{\alpha}\|\boldsymbol{\mathcal{Q}}\mu\|_0. \tag{LS1}$$

– The operators $\boldsymbol{\mathcal{Q}}\mu$ and $\mathcal{U}\mu$ are continuous. That is,

$$\|\boldsymbol{\mathcal{Q}}\mu\|_{0,e} \leq \beta\big\|A^{-1}\boldsymbol{\mathcal{Q}}\mu\big\|_{0,e} \lesssim \beta h^{-1}\|\mu\|_{h,e} \quad \text{and} \quad \|\mathcal{U}\mu\|_0 \lesssim \|\mu\|_h. \tag{LS2}$$

– The dual approximation fulfills $\boldsymbol{\mathcal{Q}}\mu \sim -\nabla_h\mathcal{U}\mu$. More precisely,

$$\begin{aligned}\|\boldsymbol{\mathcal{Q}}\mu + A\nabla_h\mathcal{U}\mu\|_0 &\lesssim \beta h^{-1}\|\mathcal{U}\mu-\mu\|_h,\\ \big\|A^{-1}\boldsymbol{\mathcal{Q}}\mu + \nabla_h\mathcal{U}\mu\big\|_0 &\lesssim h^{-1}\|\mathcal{U}\mu-\mu\|_h.\end{aligned} \tag{LS3}$$

– We have consistency with linear finite elements: if $w \in \overline{V}_h$, it holds

$$\mathcal{U}\gamma_h w = w \qquad \text{and} \qquad \boldsymbol{\mathcal{Q}}\gamma_h w = -A\nabla w. \tag{LS4}$$

– The standard spectral properties of the condensed stiffness matrix hold in the sense that

$$\alpha\|\mu\|_h^2 \lesssim a(\mu,\mu) \lesssim \beta h^{-2}\|\mu\|_h^2. \tag{LS5}$$

Remark 3.2 (Verification for the RT-H methods)**.** We have that for the RT-H method:

– (LS1) follows from tracing the constants in Lemma 3.1 of [20].
– (LS2) follows by tracing the constants in the proof of Lemma 3.3 from [18].
– (LS3) follows by a prove similar to Lemma 3.3 of [15], whose direct extension is $\|A^{-1}\boldsymbol{\mathcal{Q}}\mu + \nabla_h\mathcal{U}\mu\|_0 \lesssim h^{-1}\|\mathcal{U}\mu-\mu\|_h$.
– (LS4) can be proven directly.
– (LS5) follows by tracing the constants in the proof of Theorem 2.2 from [31].

Remark 3.3 (Other methods)**.** Proving the properties (LS1)–(LS5) for the LDG-H method relies on the respective results for the RT-H method (see Rem. 3.2). Using the arguments in [20, 41], one can recover (LS1)–(LS5) for LDG-H as well. The results also transfer to EDG since it uses the same local solvers as LDG-H.

The results also transfer to BDM-H, but we omit the explicit tracking of α and β for this method to keep the manuscript readable. Thus, in what follows, all theorems hold for BDM-H with possibly adapted dependencies on α, β.

3.4. Mesh transfer operators

To define our multiscale strategy, we need mesh transfer operators between the coarse and the fine meshes. We again emphasize that the skeletal spaces are non-nested, so particular strategies are required. More precisely, we need an injection operator $\mathcal{I}_h\colon M_H \to M_h$, that fulfills the following properties,

$$\mathcal{I}_h m_H = v|_{\mathcal{F}_h} \qquad \text{if } m_H = v|_{\mathcal{F}_H} \text{ for some } v \in \overline{V}_H, \tag{3.9a}$$

$$\|m_H\|_{H,E} \eqsim \|\mathcal{I}_h m_H\|_{h,E} \qquad \text{for } m_H \in \overline{M}_H, E \in \mathcal{E}_H. \tag{3.9b}$$

A possible choice is the operator defined in (3.1) of [41], which fulfills (3.9) by construction. As the precise definition of $\mathcal{I}_h$ is irrelevant to our analysis, we omit a detailed discussion of the construction here. Note that the authors of [41] construct several operators with remarkable stability properties. However, they do not build any stable projection operators explicitly. Unfortunately, a computable and stable projection is essential for our approach. Thus, we construct such an operator $\Pi_H\colon M_h \to M_H$ in the following.

First, we define the linear interpolation $I_H^{\mathrm{avg}}\colon V_H \to \overline{V}_H$ which uses the averaging in the vertices $\boldsymbol{x}$ of the mesh $\mathcal{E}_H$, namely

$$[I_H^{\mathrm{avg}} u](\boldsymbol{x}) = \frac{1}{N_{\boldsymbol{x}}} \sum_{i=1}^{N_{\boldsymbol{x}}} u|_{E_i}(\boldsymbol{x}).$$

Here, $N_{\boldsymbol{x}}$ is the number of elements $E_i \in \mathcal{E}_H$ meeting in the vertex $\boldsymbol{x}$ and $u|_{E_i}$ is the restriction of a function $u \in V_H$ to the element E_i, which is single-valued in $\boldsymbol{x}$. For $\boldsymbol{x} \in \partial\Omega$, we set $[I_H^{\mathrm{avg}} u](\boldsymbol{x}) = 0$. The linear interpolation I_h^{avg} is defined similarly, but the averaging is only performed within coarse elements. That is, for every $E \in \mathcal{E}_H$, we define

$$I_h^{\mathrm{avg}}|_E\colon V_h|_E \to \overline{V}_h|_E$$

by

$$[(I_h^{\mathrm{avg}}|_E)u](\boldsymbol{x}) = \frac{1}{n_{\boldsymbol{x}}} \sum_{i=1}^{n_{\boldsymbol{x}}} u|_{e_i}(\boldsymbol{x}),$$

where now $n_{\boldsymbol{x}}$ denotes the number of fine elements $e_i \in \mathcal{E}_h$ with $e_i \subset E$ that meet in the vertex $\boldsymbol{x}$. Note that I_h^{avg} is the sum of the respective contributions on coarse elements and allows for jumps across coarse element boundaries.

The linear interpolations can now be used in conjunction with the coarse element-wise L^2 projection $\mathcal{P}_H^{L^2}$ onto V_H to define a projection *via*

$$\Pi_H\colon M_h \ni \mu \mapsto \gamma_H \Big[I_H^{\mathrm{avg}} \mathcal{P}_H^{L^2} I_h^{\mathrm{avg}} \mathcal{U} \Big] \mu \in \overline{M}_H \subset M_H. \tag{3.10}$$

Clearly, $\mathcal{I}_h \Pi_H$ is a projection since $\Pi_H\colon M_h \to \overline{M}_H$, and $\mathcal{I}_h$ reproduces each function $\mu \in \overline{M}_H$ as the trace of its linear bulk extension $\gamma_h \bar{v}$ with $\bar{v} \in \overline{V}_H$ such that $\gamma_H \bar{v} = \mu$, see (3.9a). A similar argument holds for Π_H.

Lemma 3.4. *We have for all $m \in M_h$ and all $E \in \mathcal{E}_H$ that*

$$\|(\mathcal{I}_h \Pi_H) m\|_a \lesssim \sqrt{\beta/\alpha} \Big\| A^{-1/2} \boldsymbol{\mathcal{Q}} m \Big\|_0 \lesssim \sqrt{\beta/\alpha} \|m\|_a, \tag{3.11}$$

$$\|(\mathcal{I}_h \Pi_H) m\|_{h,E} \lesssim \|m\|_{h,\omega_E}. \tag{3.12}$$

Here, $\|m\|_H$ refers to the norm $\|\cdot\|_H$ of the restriction of m to the coarse skeleton.

Proof. The proof makes heavy use of (5.10) and Lemma 5.2 from [41]. First, we need a few additional estimates. Writing $[\![v]\!] = v_+ - v_-$ for the jump between two different values of v at an interface, we deduce from (5.10) and Lemma 5.2 of [41] the following inequalities.

(1) For all $v \in V_\star$, $\star \in \{h, H\}$, we have that

$$\star\|\nabla_\star I_\star^{\mathrm{avg}} v - \nabla_\star v\|_0 \lesssim \|I_\star^{\mathrm{avg}} v - v\|_0 \lesssim \star\|\nabla_\star v\|_0 + \|[\![v]\!]\|_\star. \tag{3.13}$$

In fact, only the second inequality is (5.10) of [41], while the first inequality is a standard scaling argument.

(2) For all $\mu \in M_h$, we have that

$$\|\nabla_h (I_h^{\mathrm{avg}} \mathcal{U})\mu\|_0 \lesssim \left\|A^{-1}\boldsymbol{\mathcal{Q}}\mu\right\|_0. \tag{3.14}$$

For polynomial degree 1, the upper bound in (3.13) can be established with the norm of the jump only, but the gradient term is required for higher polynomial degrees.

We now turn to show the inequality (3.11). We observe that for any $m \in M_h$

$$\begin{aligned}\|(\mathcal{I}_h\Pi_H)m\|_a &\overset{\text{Ⓐ}}{=} \|A^{-1/2}\boldsymbol{\mathcal{Q}}(\mathcal{I}_h\Pi_H)m\|_0 \overset{\text{Ⓑ}}{\simeq} \|A^{1/2}\nabla[I_H^{\mathrm{avg}}\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}]m\|_0 \\ &\overset{\text{Ⓒ}}{\lesssim} \beta^{1/2}\underbrace{\left\|\nabla_H\left[\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}\right]m\right\|_0}_{=\Xi_1} + \beta^{1/2}H^{-1}\underbrace{\left\|\left[\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}\right]m - \mathcal{U}m\right\|_H}_{=\Xi_2} + \beta^{1/2}\underbrace{H^{-1}\|\mathcal{U}m - m\|_H}_{=\Xi_3}.\end{aligned}$$

Here Ⓐ uses $(\mathcal{I}_h\Pi_H)m \in \overline{M}_h$ and the first equality in (LS4), which imply that $\gamma_h\mathcal{U}(\mathcal{I}_h\Pi_H)m = (\mathcal{I}_h\Pi_H)m$. Relation Ⓑ uses the second equality in (LS4) as well as (2.2) and (2.3). To obtain Ⓒ, we first use (3.13) (for $\star = H$) and the fact that

$$\left\|\left[\!\left[\left[\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}\right]m\right]\!\right]\right\|_H \le 2\left\|\left[\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}\right]m - m\right\|_H,$$

which in turn uses that $m = 0$ on $\partial\Omega$. Note that the fine-scale objects $\mathcal{U}m$ and m are only evaluated on the coarse skeleton in the above equations. Next, we bound the individual terms Ξ_1, Ξ_2, and Ξ_3. For Ξ_1, we have

$$\begin{aligned}\Xi_1 &\le \left\|\left(\nabla_H\left[\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}\right]m - \nabla_h[I_h^{\mathrm{avg}}\mathcal{U}]m\right)\right\|_0 + \|\nabla_h[I_h^{\mathrm{avg}}\mathcal{U}]m\|_0 \\ &\lesssim \|\nabla_h I_h^{\mathrm{avg}}\mathcal{U}m\|_0 \overset{(3.14)}{\lesssim} \left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_0 \lesssim \alpha^{-1/2}\|m\|_a,\end{aligned}$$

where the second inequality is ([21], Lem. 1.58) stating that on a single element $E \in \mathcal{E}_H$, $\|\nabla(\mathcal{P}_H^{L^2} v - v)\|_{0,E} \lesssim \|\nabla v\|_{0,E}$ for all $v \in H^1(E)$. Additionally, using the approximation properties of I_h^{avg} (see (3.13)) and $\mathcal{P}_H^{L^2}$ we estimate

$$\begin{aligned}\Xi_2 &\lesssim \left\|\left[\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}\right]m - [I_h^{\mathrm{avg}}\mathcal{U}]m\right\|_0 + \|[I_h^{\mathrm{avg}}\mathcal{U}]m - \mathcal{U}m\|_0 \\ &\lesssim H\|\nabla_h[I_h^{\mathrm{avg}}\mathcal{U}]m\|_0 + h\|\nabla_h\mathcal{U}m\|_0 + \|\mathcal{U}m - m\|_h,\end{aligned}$$

which, in turn, can be bounded by $H\alpha^{-1/2}\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\|_0$ if one applies (3.14), (LS3), and (LS1). Last, we use the definitions of $\|\cdot\|_h$ and $\|\cdot\|_H$ to write

$$\Xi_3 \le (Hh)^{-1/2}\|\mathcal{U}m - m\|_h \overset{(LS1)}{\lesssim} \sqrt{\tfrac{h}{H}}\left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_0 \overset{h<H}{\lesssim} \left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_0 \lesssim \alpha^{-1/2}\|m\|_a.$$

Altogether, this proves (3.11).

Next, we prove (3.12) by observing that

$$\|I_H^{\mathrm{avg}}\mathcal{U}m\|_{h,E} \lesssim \|\mathcal{U}m\|_{0,\omega_E} \tag{3.15}$$

for all $E \in \mathcal{E}_H$. This allows us to estimate

$$\|(\mathcal{I}_h\Pi_H)m\|_{h,E} = \left\|I_H^{\mathrm{avg}}\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}m\right\|_{0,E} \lesssim \left\|\mathcal{P}_H^{L^2} I_h^{\mathrm{avg}}\mathcal{U}m\right\|_{0,\omega_E} \le \|I_h^{\mathrm{avg}}\mathcal{U}m\|_{0,\omega_E} \lesssim \|\mathcal{U}m\|_{0,\omega_E} \lesssim \|m\|_{h,\omega_E},$$

where the first and third inequalities are (3.15), the second inequality is the stability of the L^2 projection, and the last inequality is (LS2). □

Next, we prove an interpolation-type estimate for the concatenation $\mathcal{I}_h\Pi_H$ that will be an essential ingredient to show approximation properties of the considered multiscale method.

Lemma 3.5. *We have for all $m \in M_h$ and all $E \in \mathcal{E}_H$ that*

$$\|\mathcal{U}m - \mathcal{U}\mathcal{I}_h\Pi_H m\|_{0,E} \lesssim \|m - \mathcal{I}_h\Pi_H m\|_{h,E} \lesssim H\big\|A^{-1}\boldsymbol{\mathcal{Q}}m\big\|_{0,\omega_E}.$$

For the proof of this lemma, we heavily rely on a specific lifting operator

$$S : M_h \to V_{\mathrm{disc}}^{\mathrm{c}} \cap H^1(\Omega),$$

where $V_{\mathrm{disc}}^{\mathrm{c}}$ is a finite-dimensional space of piecewise polynomials, and S fulfills

$$\mu = \pi_h\gamma_h S\mu, \tag{3.16a}$$
$$\|S\mu\|_{0,E} \eqsim \|\mu\|_{h,E}, \tag{3.16b}$$
$$\|\nabla S\mu\|_{0,E} \eqsim \big\|A^{-1}\boldsymbol{\mathcal{Q}}\mu\big\|_{0,E}. \tag{3.16c}$$

for all $\mu \in M_h$, $E \in \mathcal{E}_H$. Again, π_h is the orthogonal projection onto M_h with respect to $\langle\cdot,\cdot\rangle_h$. The operator S is rigorously defined and analyzed in Appendix A.

Proof of Lemma 3.5. First, we use (LS2) and (3.16b) to deduce

$$\|\mathcal{U}m - \mathcal{U}\mathcal{I}_h\Pi_H m\|_{0,E} \lesssim \|m - \mathcal{I}_h\Pi_H m\|_{h,E} \eqsim \|Sm - S\mathcal{I}_h\Pi_H m\|_{0,E} =: (\bigstar).$$

Notably, the equality (3.16a), the definition of Π_H in (3.10), and the property of $\mathcal{U}$ in (3.3) imply

$$(\bigstar) = \|Sm - S\mathcal{I}_h\Pi_H\pi_h\gamma_h Sm\|_{0,E} = \|Sm - S\mathcal{I}_h\Pi_H\gamma_h Sm\|_{0,E}.$$

Next, we set $z = Sm \in H^1(\Omega)$ and consider a coarse element $E \in \mathcal{E}_H$ with element patch ω_E, see (3.7). We want to use the standard scaling argument in a non-standard way. That is, we want to prove that

$$(\bigstar) = \|(\mathbf{1} - S\mathcal{I}_h\Pi_H\gamma_h)z\|_{0,E} \lesssim H^{d/2}\Big\|\widehat{\nabla}\widehat{z}\Big\|_{0,\widehat{\omega}_{\widehat{E}}} \lesssim H\|\nabla z\|_{0,\omega_E}, \tag{3.17}$$

where $\widehat{\omega}_{\widehat{E}}$ is a reference patch around the reference element $\widehat{E}$, and operators with a hat act on the reference patch. Moreover, $\widehat{z} = z \circ \Phi^{-1}$ if $\Phi\colon \widehat{\omega}_E \to \omega_E$ is a bijective, affine-linear mapping. If (3.17) holds, we can deduce the result by recalling that $z = Sm$ and $\|\nabla Sm\|_{0,E'} \lesssim \|A^{-1}\boldsymbol{\mathcal{Q}}m\|_{0,E'}$ for any $E' \in \mathcal{E}_H$ according to (3.16c).

The equality in (3.17) is evident, and the second inequality is the standard scaling argument for the H^1-seminorm of z. Therefore, it remains to show the first inequality in (3.17): we define the operator

$$\mathcal{G}_E\colon H^1\big(\widehat{\omega}_{\widehat{E}}\big) \ni \widehat{z} \mapsto ((\mathbf{1} - S\mathcal{I}_h\Pi_H\gamma_h)(\widehat{z}\circ\Phi))|_E \in L^2(E),$$

for which we want to employ the Bramble–Hilbert lemma ([16], Thm. 28.1). Obviously, we have that $\mathcal{G}_E(p) = 0$ if p is a polynomial of degree at most one. That is, we need to show that $\|\mathcal{G}_E\| \lesssim H^{d/2}$ to obtain the first inequality in (3.17). Clearly, we have that $\|\mathbf{1}(\widehat{z}\circ\Phi)\|_{0,E} \lesssim H^{d/2}\|\widehat{z}\|_{0,\widehat{E}}$. Moreover,

$$\|S\mathcal{I}_h\Pi_H\gamma_h z\|_{0,E} \overset{(3.16b)}{\eqsim} \|\mathcal{I}_h\Pi_H\gamma_h z\|_{h,E} \overset{(3.12)}{\lesssim} \|\gamma_h z\|_{h,\omega_E} \lesssim H^{d/2}\|\widehat{\gamma}_h\widehat{z}\|_{L^2(\partial\widehat{\omega}_{\widehat{E}})} \lesssim H^{d/2}\|\widehat{z}\|_{H^1(\widehat{\omega}_{\widehat{E}})} \tag{3.18}$$

by the standard scaling argument with $z = \widehat{z}\circ\Phi$. □

4. Prototypical skeletal multiscale approximation

In this section, we state and analyze a prototypical multiscale approach to discretize problem (2.1). Therefore, we introduce a multiscale space on the skeleton, which is then used as discretization space in a Galerkin fashion.

The idea of the method is built upon suitable corrections of coarse functions by finer ones. The kernel space gives an appropriate fine space for such corrections,

$$\mathcal{W}_h := \{\mu \in M_h \colon \mathcal{I}_h \Pi_H \mu = 0\} = \{\mu \in M_h \colon \Pi_H \mu = 0\}, \tag{4.1}$$

where Π_H and $\mathcal{I}_h$ are the projection and injection operator as defined in (3.10) and (3.9b), respectively. We define a so-called *correction operator* $\mathcal{C} \colon M_h \to \mathcal{W}_h$ as follows: for any $\xi \in M_h$, the function $\mathcal{C}\xi \in \mathcal{W}_h$ solves

$$a(\mathcal{C}\xi, \eta) = a(\xi, \eta) \tag{4.2}$$

for all $\eta \in \mathcal{W}_h$. Due to the coercivity of a, the operator $\mathcal{C}$ is uniquely defined and stable. Alternatively, the operator $\mathcal{C}$ can be defined based on an equivalent saddle point problem (see, *e.g.*, the derivations in [45], Ch. 2.3.2), which is beneficial mainly for implementation purposes. The alternative characterization seeks the pair $(\mathcal{C}\xi, \lambda_H) \in M_h \times \overline{M}_H$ that solves

$$\begin{aligned} a(\mathcal{C}\xi, \eta) + \langle \lambda_H, \Pi_H \eta \rangle_H &= a(\xi, \eta), \\ \langle \Pi_H \mathcal{C}\xi, \mu_H \rangle_H &= 0 \end{aligned} \tag{4.3}$$

for all $(\eta, \mu_H) \in M_h \times \overline{M}_H$. Note that this definition does not require an explicit characterization of the kernel space $\mathcal{W}_h$ (*e.g.*, in the form of a suitable basis).

Next, we define a multiscale space based on the newly defined correction operator,

$$\tilde{M}_H := (\mathbf{1} - \mathcal{C}) M_h = \mathcal{W}_h^{\perp_a}, \tag{4.4}$$

which is by definition the orthogonal complement of $\mathcal{W}_h$ with respect to a. Note that this space is actually a coarse-scale space (thus the index H) in the sense that it can be generated by the coarse space $\overline{M}_H$ and it holds that $\dim \tilde{M}_H = \dim \overline{M}_H$. To see this, we first note that $\mathcal{C}(\mathbf{1} - \mathcal{I}_h \Pi_H)\xi = (\mathbf{1} - \mathcal{I}_h \Pi_H)\xi$, because $\mathcal{C}|_{\mathcal{W}_h} = \mathbf{1}|_{\mathcal{W}_h}$. Therefore, $(\mathbf{1} - \mathcal{C})\xi = (\mathbf{1} - \mathcal{C})\mathcal{I}_h \Pi_H \xi$. Moreover, we have that

$$\tilde{M}_H = (\mathbf{1} - \mathcal{C}) M_h = (\mathbf{1} - \mathcal{C}) \mathcal{I}_h \Pi_H M_h = (\mathbf{1} - \mathcal{C}) \mathcal{I}_h \overline{M}_H$$

is generated by the coarse space $\overline{M}_H$ only. That is, $(\mathbf{1} - \mathcal{C})\mathcal{I}_h \colon \overline{M}_H \to \tilde{M}_H$ is a bijection with inverse Π_H. This follows from the definition of Π_H (see (3.10)), the properties of $\mathcal{I}_h$ (see (3.9b)), and the definition of $\mathcal{W}_h$ (see (4.1)). In the following, we abbreviate $\mathcal{R} = (\mathbf{1} - \mathcal{C})\mathcal{I}_h \colon \overline{M}_H \to M_h$. Note that, if $\{b_{H,i}\}_i$ is a (nodal) basis of the coarse space $\overline{M}_H$, then $\{\tilde{b}_{H,i}\}_i$ with

$$\tilde{b}_{H,i} := \mathcal{R} b_{H,i} \tag{4.5}$$

defines a basis of $\tilde{M}_H$. For the practical localization of the correction (cf. Sect. 6 below), an equivalent formulation based on so-called *element corrections* is beneficial. Therefore, we decompose the correction operator in (4.3) into element-wise contributions. Let $\xi_H \in \overline{M}_H$ and $E \in \mathcal{E}_H$ a coarse mesh element. The pair $(\mathcal{C}_E \xi_H, \lambda_{H,E}) \in M_h \times \overline{M}_H$ solves

$$\begin{aligned} a(\mathcal{C}_E \xi_H, \eta) + \langle \lambda_{H,E}, \Pi_H \eta \rangle_H &= a_E(\mathcal{I}_h \xi_H, \eta), \\ \langle \Pi_H \mathcal{C}_E \xi_H, \mu_H \rangle_H &= 0 \end{aligned} \tag{4.6}$$

for all $\eta \in M_h$ and $\mu_H \in \overline{M}_H$. Equivalently, the element-wise correction $\mathcal{C}_E \xi_H \in \mathcal{W}_h$ is characterized by

$$a(\mathcal{C}_E \xi_H, \eta) = a_E(\mathcal{I}_h \xi_H, \eta) \tag{4.7}$$

for all $\eta \in \mathcal{W}_h$. Here, a_E denotes the restriction of a to a coarse element $E \in \mathcal{E}_H$. By linearity arguments, we have $\mathcal{C}\mathcal{I}_h \xi_H = \sum_{E \in \mathcal{E}_H} \mathcal{C}_E \xi_H$.

The non-localized prototypical multiscale method now seeks $\tilde{m}_H \in \tilde{M}_H$ such that

$$a(\tilde{m}_H, \tilde{\mu}_H) = \int_\Omega f\,\mathcal{U}\tilde{\mu}_H \,\mathrm{d}x \qquad \text{for all } \tilde{\mu}_H \in \tilde{M}_H. \tag{4.8}$$

Note that we have stability of $\tilde{m}_H$ in the sense that

$$\|\tilde{m}_H\|_a \lesssim \alpha^{-1/2}\|f\|_0 \tag{4.9}$$

due to (LS2), (LS5), and using $\tilde{\mu}_H = \tilde{m}_H$ in (4.8).

5. Error analysis for the prototypical method

The following theorem quantifies the error of the prototypical multiscale method in terms of the coarse mesh parameter H.

Theorem 5.1 (Error of the prototypical method)**.** *Let $f \in L^2(\Omega)$ and let $m \in M_h$ be the solution of (3.4). Then*

$$\alpha^{1/2}\|\mathcal{U}m - \mathcal{U}\tilde{m}_H\|_0 + \left\|A^{-1/2}(\boldsymbol{\mathcal{Q}}m - \boldsymbol{\mathcal{Q}}\tilde{m}_H)\right\|_0 \lesssim \|m - \tilde{m}_H\|_a \lesssim \alpha^{-1/2}H\,\|f\|_0. \tag{5.1}$$

Remark 5.2. The operator $\mathcal{I}_h\Pi_H$ produces the same result when applied to the solutions m and $\tilde{m}_H$, respectively. In that sense, the *coarse-scale contributions* of the two functions are identical. This is an important property to prove Theorem 5.1. Indeed, by the construction of the correction operator in (4.2) and the multiscale space in (4.4), we have that

$$a(\tilde{\mu}_H, \eta) = 0$$

for all $\tilde{\mu}_H \in \tilde{M}_H$ and $\eta \in \mathcal{W}_h$. In particular, the above property uniquely determines the space $\mathcal{W}_h$. Since Galerkin orthogonality implies that

$$a(m - \tilde{m}_H, \tilde{\mu}_H) = 0$$

for all $\tilde{\mu}_H \in \tilde{M}_H$, we deduce from the symmetry of a that $m - \tilde{m}_H \in \mathcal{W}_h$ and, thus, $\mathcal{I}_h\Pi_H(m - \tilde{m}_H) = 0$.

Proof of Theorem 5.1. The first inequality of (5.1) follows from the linearity of $\mathcal{U}$ and $\boldsymbol{\mathcal{Q}}$,

$$\|\mathcal{U}m - \mathcal{U}\tilde{m}_H\|_0 = \|\mathcal{U}(m - \tilde{m}_H)\|_0 \overset{\text{(LS2)}}{\lesssim} \|m - \tilde{m}_H\|_h \overset{\text{(LS5)}}{\lesssim} \alpha^{-1/2}\|m - \tilde{m}_H\|_a \tag{5.2}$$

and the the definition of $\|\cdot\|_a$, *i.e.*,

$$\left\|A^{-1/2}(\boldsymbol{\mathcal{Q}}m - \boldsymbol{\mathcal{Q}}\tilde{m}_H)\right\|_0^2 = \left\|A^{-1/2}\boldsymbol{\mathcal{Q}}(m - \tilde{m}_H)\right\|_0^2 = \int_\Omega A^{-1}[\boldsymbol{\mathcal{Q}}(m - \tilde{m}_H)]^2 \,\mathrm{d}x \le \|m - \tilde{m}_H\|_a^2.$$

Since $\tilde{m}_H \in \tilde{M}_H \subset M_h$, it is a valid test function for both (4.8) and (3.4). This implies that $a(m - \tilde{m}_H, \tilde{m}_H) = 0$. This and the symmetry of a can be used to deduce that

$$\begin{aligned}
\|m - \tilde{m}_H\|_a^2 &= a(m - \tilde{m}_H, m - \tilde{m}_H) = a(m, m - \tilde{m}_H) \\
&= \int_\Omega f\,\mathcal{U}(m - \tilde{m}_H)\,\mathrm{d}x \le \|f\|_0\,\|\mathcal{U}(m - \tilde{m}_H)\|_0 \\
&= \|f\|_0\,\|\mathcal{U}(m - \tilde{m}_H) - \mathcal{U}\mathcal{I}_h\Pi_H(m - \tilde{m}_H)\|_0 \\
&\lesssim \|f\|_0\,\alpha^{-1/2}H\,\|m - \tilde{m}_H\|_a,
\end{aligned} \tag{5.3}$$

where we use that m solves (3.4) to get the third equality. The fourth equality uses that $\mathcal{I}_h\Pi_H(m - \tilde{m}_H) = 0$ (cf. Rem. 5.2), and the final inequality uses Lemma 3.5 and

$$\left\|A^{-1}\boldsymbol{\mathcal{Q}}(m - \tilde{m}_H)\right\|_0 \le \alpha^{-1/2}\|m - \tilde{m}_H\|_a.$$

Dividing both sides of the inequality (5.3) by $\|m - \tilde{m}_H\|_a$, we obtain the final result. □

The previous theorem states that the computed multiscale solution (with a moderate number of degrees of freedom) is a good approximation of the fine-scale HDG solution computed in (3.4). However, the computation of the correction operator $\mathcal{C}$ or, in a more practical manner, of the basis functions $\tilde{b}_{H,i}$ defined in (4.5) requires global fine-scale problems to be solved. This is not feasible in practice, so the localization of these problems is an important task that will be studied in the following. First, we show the following useful result.

Lemma 5.3 (Local pre-image)**.** *Let $n \in \mathbb{N}$ and for each $i = 1, \dots, n$ let $\omega_i \subset \Omega$ be a convex union of (coarse) elements of $\mathcal{E}_H$ with $\operatorname{diam} \omega_i \sim H$ and*

$$\omega := \operatorname{int}\left(\bigcup_{i=1}^{n} \omega_i \right).$$

For each i, we denote with $\Gamma_i \subset \partial\omega_i \cap \partial\omega$ a part of the boundary with non-zero measure. Further, let $\xi \in M_h$ with $\operatorname{supp}(\Pi_H \xi) \subset \omega$ and $\xi|_{\Gamma_i} = 0$ for all $i = 1, \dots, n$. Then there exists a function $\mathfrak{b} \in M_h(\omega)$ with

$$\Pi_H \mathfrak{b} = \Pi_H \xi \quad \text{and} \quad \|\mathfrak{b}\|_{a,\omega} \lesssim (\beta/\alpha)^{1/2} \left\| A^{-1/2} \boldsymbol{\mathcal{Q}} \xi \right\|_{0,\omega}.$$

Proof. We define $\mathfrak{b} \in M_h(\omega)$ (together with the Lagrange multiplier $\delta_H \in \overline{M}_H(\omega)$ as the solution of

$$\begin{aligned} a(\mathfrak{b}, \eta) - \langle \delta_H, \Pi_H \eta \rangle_H &= 0, \\ \langle \Pi_H \mathfrak{b}, \mu_H \rangle_H &= \langle \Pi_H \xi, \mu_H \rangle_H \end{aligned} \tag{5.4}$$

for all $\eta \in M_h(\omega)$ and all $\mu_H \in \overline{M}_H(\omega)$. Note that $\mathfrak{b}$ and δ_H can be trivially extended by zero to the spaces M_h and $\overline{M}_H$, respectively. From classical saddle point theory (see, *e.g.*, [12], Cor. 4.2.1), we know that (5.4) has a unique solution if the inf-sup condition

$$\inf_{\mu_H \in \overline{M}_H(\omega)} \sup_{\zeta \in M_h(\omega)} \frac{\langle \mu_H, \Pi_H \zeta \rangle_H}{\|\mu_H\|_H \, \|\zeta\|_a} \geq \gamma(H) > 0 \tag{5.5}$$

holds and a is coercive.

To show the inf-sup condition (5.5), let $\mu_H \in \overline{M}_H(\omega)$. Then, with the explicit choice $\zeta = \mathcal{I}_h \mu_H \in \overline{M}_h(\omega)$ we have that $\mu_H = \Pi_H \mathcal{I}_h \mu_H$, and we can deduce that

$$\sup_{\zeta \in M_h(\omega)} \frac{\langle \mu_H, \Pi_H \zeta \rangle_H}{\|\mu_H\|_H \, \|\zeta\|_a} \geq \frac{\langle \mu_H, \mu_H \rangle_H}{\|\mu_H\|_H \, \|\mathcal{I}_h \mu_H\|_a} \overset{(3.9b)}{\eqsim} \frac{\|\mathcal{I}_h \mu_H\|_h}{\|\mathcal{I}_h \mu_H\|_a}. \tag{5.6}$$

Next, let $v \in \overline{V}_H$ be such that $\gamma_H v = \mu_H$. This implies that $\gamma_h v = \mathcal{I}_h \mu_H$. By (LS4) and the definition of a we have that

$$\|\mathcal{I}_h \mu_H\|_a \lesssim \beta^{1/2} \|\nabla v\|_0 =: (\bigstar).$$

Now, by the standard scaling argument on the coarse mesh, we have

$$(\bigstar) \lesssim \beta^{1/2} H^{-1} \|v\|_0 \lesssim \beta^{1/2} H^{-1} \|\mathcal{I}_h \mu_H\|_h,$$

where the second estimate follows from the equivalence of $\|\cdot\|_0$ and $\|\cdot\|_h$ on $\overline{V}_h$. Going back to (5.6), we obtain the inf-sup condition (5.5) with $\gamma(H) \gtrsim \beta^{-1/2} H > 0$.

Due to the locality of $\mathfrak{b}$ and $\Pi_H \xi$ and with Corrollary 4.2.1 of [12] and (5.5), problem (5.4) is well-posed and the stability estimate

$$\|\mathfrak{b}\|_{a,\omega} \leq \frac{2}{\gamma(H)} \|\Pi_H \xi\|_{H,\omega}$$

holds. Using (3.9b), we obtain

$$\|\mathfrak{b}\|_{a,\omega} \lesssim \beta^{1/2}H^{-1}\,\|\Pi_H\xi\|_{H,\omega} \lesssim \beta^{1/2}H^{-1}\|\mathcal{I}_h\Pi_H\xi\|_{h,\omega} \lesssim \beta^{1/2}H^{-1}\|\mathcal{I}_h\Pi_H\xi-\xi\|_{h,\omega} + \beta^{1/2}H^{-1}\|\xi\|_{h,\omega}.$$

The first term can be estimated using Lemma 3.5. For the second term, we require a Poincaré–Friedrichs inequality for broken skeletal spaces. This result can be derived from Theorem B.3 in the appendix, which follows from [13] when carefully tracking the dependence on the domain size. Overall, we obtain

$$\|\mathfrak{b}\|_{a,\omega} \lesssim (\beta/\alpha)^{1/2}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}\xi\right\|_{0,\omega}.$$

Finally, we emphasize that $\Pi_H\mathfrak{b} = \Pi_H\xi$ by construction. □

Theorem 5.4 (Decay of element corrections)**.** *Let $\xi_H \in M_H$ and $E \in \mathcal{E}_H$ a coarse element. Further, let $\ell \in \mathbb{N}$. Then there exists $0 < \vartheta < 1$ depending on α and β such that*

$$\|\mathcal{C}_E\xi_H\|_{a,\Omega\setminus\mathsf{N}^\ell(E)} \lesssim \vartheta^\ell\|\mathcal{C}_E\xi_H\|_a. \tag{5.7}$$

To prove this essential assertion, we introduce some auxiliary results. In particular, we require a cutoff function η, which for given $\ell \in \mathbb{N}$, $\ell \geq 4$, and a coarse element $E \in \mathcal{E}_H$ is the uniquely defined continuous function that satisfies

$$\eta \equiv 0 \qquad \text{in } \mathsf{N}^{\ell-3}(E), \tag{5.8a}$$
$$\eta \equiv 1 \qquad \text{in } \Omega\setminus\mathsf{N}^{\ell-2}(E), \tag{5.8b}$$
$$\eta \text{ is linear} \qquad \text{in } E' \in \mathcal{E}_H \text{ with } \operatorname{int} E' \subset R := \mathsf{N}^{\ell-2}(E)\setminus\mathsf{N}^{\ell-3}(E), \tag{5.8c}$$

using the definition of patches in (3.8). Additionally, we define the element-wise constant function $\bar\eta$ for any $e \in \mathcal{E}_h$ by

$$\bar\eta|_e \equiv \frac{1}{|\partial e|}\int_{\partial e}\eta\,\mathrm{d}\sigma. \tag{5.9}$$

Further, we denote again with π_h the face-wise L^2 projection onto M_h. The following results are important in the subdomains where η is not constant. They trivially hold in parts of the domain where η is constant.

Lemma 5.5. *Let $E \in \mathcal{E}_H$ a coarse mesh cell, η be defined as in (5.8) for some $\ell \in \mathbb{N}$. Then, for $E' \in \mathcal{E}_H$ and $\mathcal{E}_{E'} = \{e \in \mathcal{E}_h : e \subset E'\}$*

$$\sum_{e\in\mathcal{E}_{E'}}\int_e A^{-1}\boldsymbol{\mathcal{Q}}m\cdot[\boldsymbol{\mathcal{Q}}(m\eta)-\eta\boldsymbol{\mathcal{Q}}m]\,\mathrm{d}x \lesssim \frac{\max\{1,\beta\}}{\min\{1,\alpha\}}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|^2_{0,\omega_{E'}}.$$

for any $m \in M_h$ with $(\mathcal{I}_h\Pi_H m)|_{\mathcal{E}_{E'}} = 0$.

Proof. We start with estimating the term

$$\int_e A^{-1}\boldsymbol{\mathcal{Q}}m\cdot[\boldsymbol{\mathcal{Q}}(m\eta)-\eta\boldsymbol{\mathcal{Q}}m]\,\mathrm{d}x = \underbrace{\int_e A^{-1}\boldsymbol{\mathcal{Q}}m\cdot[\boldsymbol{\mathcal{Q}}(m\eta)-\boldsymbol{\mathcal{Q}}(m\bar\eta)]\,\mathrm{d}x}_{=:\Xi_1} + \underbrace{\int_e A^{-1}\boldsymbol{\mathcal{Q}}m\cdot[\bar\eta\boldsymbol{\mathcal{Q}}m-\eta\boldsymbol{\mathcal{Q}}m]\,\mathrm{d}x}_{=:\Xi_2}.$$

for any $e \in \mathcal{E}_{E'}$. Note that we have used that $\bar\eta$ is constant and thus $\boldsymbol{\mathcal{Q}}(m\bar\eta) = (\boldsymbol{\mathcal{Q}}m)\bar\eta$. Next, we employ Lemma 3.1 with $\rho = m$ and $\xi = m\eta$ as well as $\xi = m\bar\eta$ to reformulate

$$\Xi_1 = \int_{\partial e}(\boldsymbol{\mathcal{Q}}m\cdot\boldsymbol{\nu})m(\eta-\bar\eta)\,\mathrm{d}\sigma \tag{T1}$$

$$+\int_{\partial e}\tau(\mathcal{U}m-m)[\mathcal{U}(m\eta)-\mathcal{U}(m\bar{\eta})]\,\mathrm{d}\sigma, \tag{T2}$$

whose components we can estimate separately. First, we observe that

$$|(\mathrm{T1})|\lesssim h^{-1}\|\boldsymbol{\mathcal{Q}}m\|_{0,e}\|m\|_{h,e}\|\eta-\bar{\eta}\|_{L^\infty(\partial e)}\lesssim\beta^{1/2}H^{-1}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|m\|_{h,e},$$

since we can bound $\|\eta-\bar{\eta}\|_{L^\infty(\partial e)}\lesssim h/H$ (which in turn follows directly from (5.8) and (5.9)). Second, we have

$$\begin{aligned}
|(\mathrm{T2})|&\lesssim\tau_{\max}h^{-1}\|\mathcal{U}m-m\|_{h,e}\|\mathcal{U}(m\eta-m\bar{\eta})\|_{h,e}\\
&\lesssim\tau_{\max}\left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|\pi_h(m\eta)-m\bar{\eta}\|_{h,e}=\tau_{\max}\left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|\pi_h(m\eta-m\bar{\eta})\|_{h,e}\\
&\le\tau_{\max}\left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|m\eta-m\bar{\eta}\|_{h,e}\lesssim\tau_{\max}\left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|m\|_{h,e}\|\eta-\bar{\eta}\|_{L^\infty(\partial e)}\\
&\lesssim\frac{\tau_{\max}h}{\alpha^{1/2}H}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|m\|_{h,e},
\end{aligned}$$

where $\tau_{\max}=\max\{\tau(x)\colon x\in\partial e\}$ and the second inequality uses (3.3), (LS1), and (LS2). If $\tau_{\max}h\lesssim 1$, this implies that

$$|\Xi_1|\lesssim\frac{\max\{1,\beta\}}{\min\{1,\alpha\}}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}^2+H^{-2}\|m\|_{h,e}^2.$$

We also have that

$$|\Xi_2|\le\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}^2\|\bar{\eta}-\eta\|_{L^\infty(e)}\lesssim\frac{h}{H}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}^2.$$

Finally, since $\mathcal{I}_h\Pi_H m=0$ we observe that with Lemma 3.5

$$\sum_{e\subset E'}\|m\|_{h,e}^2=\|m\|_{h,E'}^2=\|m-\mathcal{I}_h\Pi_H m\|_{h,E'}^2\lesssim H^2\left\|A^{-1}\boldsymbol{\mathcal{Q}}m\right\|_{0,\omega_{E'}}^2\lesssim\alpha^{-1}H^2\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,\omega_{E'}}^2,$$

which completes the proof. □

Lemma 5.6. *Let η be defined as in (5.8). Then, for $E'\in\mathcal{E}_H$ and $\mathcal{E}_{E'}=\{e\in\mathcal{E}_h\colon e\subset E'\}$ we have*

$$\sum_{e\in\mathcal{E}_{E'}}\int_{\partial e}\tau(\mathcal{U}m-m)[\mathcal{U}(m\eta)-\eta\mathcal{U}m]\,\mathrm{d}\sigma\lesssim\alpha^{-1}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,\omega_{E'}}^2$$

for any $m\in M_h$ with $(\mathcal{I}_h\Pi_H m)|_{\mathcal{E}_{E'}}=0$.

Proof. As above, we start to rewrite the term for a single element $e\in\mathcal{E}_{E'}$,

$$\int_{\partial e}\tau(\mathcal{U}m-m)[\mathcal{U}(m\eta)-\eta\mathcal{U}m]\,\mathrm{d}\sigma=\underbrace{\int_{\partial e}\tau(\mathcal{U}m-m)[\mathcal{U}(m\eta)-\mathcal{U}(m\bar{\eta})]\,\mathrm{d}\sigma}_{=\Psi_1}+\underbrace{\int_{\partial e}\tau(\mathcal{U}m-m)[\bar{\eta}\mathcal{U}m-\eta\mathcal{U}m]\,\mathrm{d}\sigma}_{=\Psi_2}.$$

The estimate for Ψ_1 is the estimate of (T2). For the latter term, we can proceed using the same arguments as for (T2) and obtain

$$|\Psi_2|\lesssim\frac{\tau_{\max}h}{\alpha^{1/2}H}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|\mathcal{U}m\|_{h,e}\lesssim\frac{1}{\alpha^{1/2}H}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}\|m\|_{h,e}.$$

The second inequality follows directly from (LS2) and the fact that $\|q\|_{h,e}\lesssim\|q\|_{0,e}$ for any polynomial q. The remainder of the proof is based on the same arguments as in the proof of Lemma 5.5. □

Lemma 5.7. *Let η be as defined in (5.8). Then, for $E' \in \mathcal{E}_H$*

$$\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}(m\eta)\right\|_{0,E'}^2 \lesssim \frac{\max\{1,\beta\}}{\min\{1,\alpha\}}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,\omega_{E'}}^2$$

for any $m \in M_h$ with $(\mathcal{I}_h\Pi_H m)|_{\mathcal{E}_{E'}} = 0$.

Proof. Let as before $e \in \mathcal{E}_{E'} = \{e \in \mathcal{E}_h : e \subset E'\}$. Then

$$\begin{aligned}&\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}(m\eta)\right\|_{0,e}^2\\ &\quad= \underbrace{\int_e A^{-1}\boldsymbol{\mathcal{Q}}(m\eta)\cdot[\boldsymbol{\mathcal{Q}}(m\eta)-\boldsymbol{\mathcal{Q}}(m\bar\eta)]\,\mathrm{d}x}_{=:\Phi_1} + \underbrace{\int_e A^{-1}\boldsymbol{\mathcal{Q}}(m\eta)\cdot[\bar\eta\boldsymbol{\mathcal{Q}}m-\eta\boldsymbol{\mathcal{Q}}m]\,\mathrm{d}x}_{=:\Phi_2} + \underbrace{\int_e A^{-1}\boldsymbol{\mathcal{Q}}(m\eta)\cdot(\boldsymbol{\mathcal{Q}}m)\eta\,\mathrm{d}x}_{=:\Phi_3}.\end{aligned}$$

The terms Φ_1 and Φ_2 can be estimated analogously to Lemma 5.5. In particular, we use Lemma 3.1 with $\rho = \eta$ and $\xi = \eta$ as well as $\xi = \bar\eta$. The term Φ_3 can be bounded by

$$|\Phi_3| \le \left\|A^{-1/2}\boldsymbol{\mathcal{Q}}(m\eta)\right\|_{0,e}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,e}.$$

Combined with the estimates for Φ_1 and Φ_2 and summing over all fine elements in E', we obtain the assertion. □

Proof of Theorem 5.4. Let $m = \mathcal{C}_E\xi_H \in M_h$, $\ell \in \mathbb{N}$ with $\ell \ge 4$, and η as defined in (5.8) for a given element $E \in \mathcal{E}_H$. Using again the patch definition in (3.8), we have

$$\begin{aligned}\|m\|_{a,\Omega\setminus\mathsf{N}^\ell(E)}^2 &= \int_{\Omega\setminus\mathsf{N}^\ell(E)} A^{-1}\boldsymbol{\mathcal{Q}}m\cdot(\boldsymbol{\mathcal{Q}}m)\eta\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h\,:\,e\subset E}\tau\int_{\partial e}(\mathcal{U}m-m)(\mathcal{U}m-m)\eta\,\mathrm{d}\sigma\\ &\le \int_\Omega A^{-1}\boldsymbol{\mathcal{Q}}m\cdot(\boldsymbol{\mathcal{Q}}m)\eta\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h}\tau\int_{\partial e}(\mathcal{U}m-m)(\mathcal{U}m-m)\eta\,\mathrm{d}\sigma\\ &= \int_\Omega A^{-1}\boldsymbol{\mathcal{Q}}m\cdot\boldsymbol{\mathcal{Q}}(m\eta)\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h}\tau\int_{\partial e}(\mathcal{U}m-m)(\mathcal{U}(m\eta)-m\eta)\,\mathrm{d}\sigma\\ &\quad+\int_R A^{-1}\boldsymbol{\mathcal{Q}}m\cdot[(\boldsymbol{\mathcal{Q}}m)\eta-\boldsymbol{\mathcal{Q}}(m\eta)]\,\mathrm{d}x + \sum_{e\in\mathcal{E}_h\,:\,e\subset\overline{R}}\tau\int_{\partial e}(\mathcal{U}m-m)[(\mathcal{U}m)\eta-\mathcal{U}(m\eta)]\,\mathrm{d}\sigma,\end{aligned}$$

where $R = \mathsf{N}^{\ell-2}(E)\setminus\mathsf{N}^{\ell-3}(E)$. Note that $(\mathcal{I}_h\Pi_H m)|_{\mathsf{N}^1(R)} = 0$ by construction of $\mathcal{C}_E$. For the second to last and last line, we can thus use Lemmas 5.5 and 5.6, respectively, as well as

$$\sum_{E'\in\mathcal{E}_H\,:\,E'\subset\overline{\mathsf{N}^1(R)}}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,E'}^2 \le \|m\|_{a,\mathsf{N}^1(R)}^2, \tag{5.10}$$

which yields

$$\|m\|_{a,\Omega\setminus\mathsf{N}^\ell(E)}^2 \lesssim a(m,\pi_h(m\eta)) + \frac{\max\{1,\beta\}}{\min\{1,\alpha\}}\|m\|_{a,\mathsf{N}^1(R)}^2, \tag{5.11}$$

where π_h is the face-wise L^2-orthogonal projection onto M_h as above. To bound the first term, we observe that $\operatorname{supp}(\Pi_H\pi_h(m\eta)) \subset \mathsf{N}^1(R)$ and $m|_{\partial(\mathsf{N}^{\ell-4}(E))} = 0$. Therefore, by Lemma 5.3, there exists a function $\mathfrak{b} \in M_h(\mathsf{N}^1(R))$ with

$$\|\mathfrak{b}\|_{a,\mathsf{N}^1(R)} \lesssim (\beta/\alpha)^{1/2}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}\pi_h(m\eta)\right\|_{0,\mathsf{N}^1(R)} \quad\text{and}\quad \Pi_H(m\eta) = \Pi_H\mathfrak{b}.$$

We further have

$$a(m,\pi_h(m\eta)) = a(m,\pi_h(m\eta)-\mathfrak{b}) + a(m,\mathfrak{b}) \lesssim (\beta/\alpha)^{1/2}\|m\|_{a,\mathsf{N}^1(R)}\|A^{-1/2}\boldsymbol{\mathcal{Q}}\pi_h(m\eta)\|_{0,\mathsf{N}^1(R)},$$

using that $\Pi_H(\pi_h(m\eta)-\mathfrak{b})=0$ and thus

$$a(m,\pi_h(m\eta)-\mathfrak{b}) = a_E(\mathcal{I}_h\xi_H,\pi_h(m\eta)-\mathfrak{b}) = 0$$

by the first line of (4.6) and the fact that $\operatorname{supp}(\pi_h(m\eta)-\mathfrak{b})\cap E=\emptyset$. By (3.6) and (LS1) and $\boldsymbol{\mathcal{Q}}(m\eta) = \boldsymbol{\mathcal{Q}}(\pi_h(m\eta))$, we have that

$$\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}\pi_h(m\eta)\right\|_{0,\mathsf{N}^1(R)} = \left\|A^{-1/2}\boldsymbol{\mathcal{Q}}(m\eta)\right\|_{0,\mathsf{N}^{\ell-1}(E)\setminus\mathsf{N}^{\ell-3}(E)} \lesssim \left(\frac{\max\{1,\beta\}}{\min\{1,\alpha\}}\right)^{1/2}\left\|A^{-1/2}\boldsymbol{\mathcal{Q}}m\right\|_{0,\mathsf{N}^{\ell}(E)\setminus\mathsf{N}^{\ell-4}(E)},$$

where we use Lemma 5.7 for all $E'\in\mathcal{E}_H$ with $\operatorname{int}E'\subset\mathsf{N}^{\ell-1}(E)\setminus\mathsf{N}^{\ell-3}(E)$ in the last step. Going back to (5.11) and using once more (5.10), we conclude that there exists a constant $C_{\alpha,\beta}>0$ scaling like $\max\{1,\beta\}/\min\{1,\alpha\}$ with

$$\|m\|^2_{a,\Omega\setminus\mathsf{N}^\ell(E)} \le C_{\alpha,\beta}\,\|m\|^2_{a,\mathsf{N}^\ell(E)\setminus\mathsf{N}^{\ell-4}(E)} \le C_{\alpha,\beta}\,\|m\|^2_{a,\Omega\setminus\mathsf{N}^{\ell-4}(E)} - C_{\alpha,\beta}\,\|m\|^2_{a,\Omega\setminus\mathsf{N}^\ell(E)}.$$

Therefore, we obtain

$$\|m\|^2_{a,\Omega\setminus\mathsf{N}^\ell(E)} \le \frac{C_{\alpha,\beta}}{C_{\alpha,\beta}+1}\|m\|^2_{a,\Omega\setminus\mathsf{N}^{\ell-4}(E)} \le \left(\frac{C_{\alpha,\beta}}{C_{\alpha,\beta}+1}\right)^{\lfloor\ell/4\rfloor}\|m\|^2_a.$$

That is, the estimate (5.7) holds with $\vartheta = \big(C_{\alpha,\beta}/(C_{\alpha,\beta}+1)\big)^{1/8}$. The results hold for $\ell\le 3$ by slightly adjusting the constant. □

6. A practical localized multiscale method

As mentioned before, a major drawback of the above multiscale approach is that it evolves around solving auxiliary corrections globally. Based on the decaying behavior of these corrections, which was proved in the previous section, we now introduce a localized – and thus practical – method that achieves similar convergence properties as its non-localized prototypical version.

Following the ideas of the localization strategy for the LOD in the conforming setting as introduced in [36], we define an element-wise localization procedure. The localization of the element corrections defined in (4.6) leads to the following saddle point problem: find $(\mathcal{C}^\ell_E\xi_H,\lambda^\ell_{H,E})\in M_h(\mathsf{N}^\ell(E))\times\overline{M}_H(\mathsf{N}^\ell(E))$ such that

$$\begin{aligned} a\big(\mathcal{C}^\ell_E\xi_H,\eta\big) + \big\langle\lambda^\ell_{H,E},\Pi_H\eta\big\rangle_H &= a_E(\mathcal{I}_h\xi_H,\eta),\\ \big\langle\Pi_H\mathcal{C}^\ell_E\xi_H,\mu_H\big\rangle_H &= 0, \end{aligned} \tag{6.1}$$

for all $(\eta,\mu_H)\in M_h(\mathsf{N}^\ell(E))\times\overline{M}_H(\mathsf{N}^\ell(E))$. We then define the operator $\mathcal{C}^\ell$ as the sum of the respective contributions, *i.e.*, $\mathcal{C}^\ell\mathcal{I}_h\xi_H = \sum_{E\in\mathcal{E}_H}\mathcal{C}^\ell_E\xi_H$ for any $\xi_H\in\overline{M}_H$. The localized version of $\mathcal{R}$ is now given as $\mathcal{R}^\ell\colon\overline{M}_H\to M_h$ with $\mathcal{R}^\ell := (\mathbf{1}-\mathcal{C}^\ell)\mathcal{I}_h$. A corresponding localized basis is given by $\{\tilde{b}^\ell_{H,i}\}_i$ with $\tilde{b}^\ell_{H,i} := \mathcal{R}^\ell b_{H,i}$, where $\{b_{H,i}\}_i$ is once again a nodal basis of $\overline{M}_H$.

We can now define a localized version of the method defined in (4.8) that seeks $\tilde{m}^\ell_H\in\tilde{M}^\ell_H := \mathcal{R}^\ell\overline{M}_H$ such that

$$a\big(\tilde{m}^\ell_H,\tilde{\mu}^\ell_H\big) = \int_\Omega f\,\mathcal{U}\tilde{\mu}^\ell_H\,\mathrm{d}x \qquad \text{for all } \tilde{\mu}^\ell_H\in\tilde{M}^\ell_H. \tag{6.2}$$

We have the following theorem.

Theorem 6.1 (Error of the localized multiscale method)**.** *For the localized multiscale solution $\tilde{m}_H^\ell$ defined in (6.2), it holds that*

$$\|m - \tilde{m}_H^\ell\|_a \lesssim \alpha^{-1/2} H \|f\|_0 + \frac{\max\{1,\beta^{3/2}\}}{\min\{1,\alpha^2\}} \theta^\ell \|f\|_0$$

with $0 < \theta < 1$ as in Theorem 5.4 and the HDG solution $m \in M_h$ to (3.4).

Moreover, for $\ell \gtrsim \log \frac{1}{H}$, we retain the linear convergence rate of the ideal method as quantified in Theorem 5.1.

Proof. Let $\ell \geq 3$. By definition, $\tilde{m}_H^\ell$ is the best approximation of m in the space $\tilde{M}_H^\ell$ with respect to the norm $\|\bullet\|_a$. Therefore, we get with the triangle inequality, $\mathcal{R}^\ell \Pi_H \tilde{m}_H \in \tilde{M}_H^\ell$, and the equality $\mathcal{R}\Pi_H \tilde{m}_H = \tilde{m}_H$

$$\|m - \tilde{m}_H^\ell\|_a \leq \|m - \mathcal{R}^\ell \Pi_H \tilde{m}_H\|_a \leq \|m - \tilde{m}_H\|_a + \|(\mathcal{R} - \mathcal{R}^\ell)\Pi_H \tilde{m}_H\|_a. \tag{6.3}$$

Recall that $\tilde{m}_H$ is the prototypical multiscale solution to (4.8). The first term can be estimated using Theorem 5.1; for the second term, we want to use Theorem 5.4. First, we write

$$\|(\mathcal{R} - \mathcal{R}^\ell)\Pi_H \tilde{m}_H\|_a \leq \sum_{E \in \mathcal{E}_H} \|(\mathcal{C}_E - \mathcal{C}_E^\ell)\Pi_H \tilde{m}_H\|_a \tag{6.4}$$

with the operators $\mathcal{C}_E$ and $\mathcal{C}_E^\ell$ defined in (4.6) and (6.1), respectively. To further estimate the right-hand side, we once again require a cutoff function χ, which for given $\ell \in \mathbb{N}$, $\ell \geq 3$ and a coarse element $E \in \mathcal{E}_H$ is the uniquely defined continuous function that satisfies

$$\chi \equiv 0 \qquad \text{in } \mathsf{N}^{\ell-2}(E), \tag{6.5a}$$

$$\chi \equiv 1 \qquad \text{in } \Omega \setminus \mathsf{N}^{\ell-1}(E), \tag{6.5b}$$

$$\chi \text{ is linear} \qquad \text{in } E' \in \mathcal{E}_H \text{ with } \operatorname{int} E' \subset R := \mathsf{N}^{\ell-1}(E) \setminus \mathsf{N}^{\ell-2}(E), \tag{6.5c}$$

using the element patches defined in (3.8). We abbreviate $\xi_H := \Pi_H \tilde{m}_H$. Since $(\mathbf{1} - \mathcal{I}_h \Pi_H)\chi \mathcal{C}_E \xi_H \in M_h(\mathsf{N}^\ell(E)) \cap \mathcal{W}_h$ and $\mathcal{C}_E^\ell \xi_H \in M_h(\mathsf{N}^\ell(E)) \cap \mathcal{W}_h$, we obtain from the definition of $\mathcal{C}_E$ and $\mathcal{C}_E^\ell$

$$\begin{aligned}\|(\mathcal{C}_E - \mathcal{C}_E^\ell)\xi_H\|_a^2 &= a\big((\mathcal{C}_E - \mathcal{C}_E^\ell)\xi_H, (\mathcal{C}_E - \mathcal{C}_E^\ell)\xi_H\big) = a\big((\mathcal{C}_E - \mathcal{C}_E^\ell)\xi_H, \mathcal{C}_E \xi_H\big) \\ &= a\big((\mathcal{C}_E - \mathcal{C}_E^\ell)\xi_H, (\mathbf{1} - \mathcal{I}_h \Pi_H)(1 - \chi)\mathcal{C}_E \xi_H\big).\end{aligned}$$

Using Lemma 3.4 (particularly the inequalities (3.11)), Lemma 5.7 (with $m = \mathcal{C}_E \xi_H$, $\eta = 1 - \chi$) and (LS1) locally, we get

$$\|(\mathcal{C}_E - \mathcal{C}_E^\ell)\xi_H\|_a \lesssim \frac{\max\{1,\beta\}}{\min\{1,\alpha\}} \|\mathcal{C}_E \xi_H\|_{a, \mathsf{N}^\ell(E) \setminus \mathsf{N}^{\ell-3}(E)} \lesssim \frac{\max\{1,\beta\}}{\min\{1,\alpha\}} \|\mathcal{C}_E \xi_H\|_{a, \Omega \setminus \mathsf{N}^{\ell-3}(E)}.$$

Going back to (6.4) and using Theorem 5.4, we arrive at

$$\|(\mathcal{R} - \mathcal{R}^\ell)\xi_H\|_a \leq \sum_{E \in \mathcal{E}_H} \|(\mathcal{C}_E - \mathcal{C}_E^\ell)\xi_H\|_a \lesssim \frac{\max\{1,\beta\}}{\min\{1,\alpha\}} \sum_{E \in \mathcal{E}_H} \theta^{\ell-3} \|\mathcal{C}_E \xi_H\|_a.$$

As a last step, we use the stability of $\mathcal{C}_E$, *i.e.*,

$$\|\mathcal{C}_E \xi_H\|_a \lesssim \|\mathcal{I}_h \xi_H\|_{a,E} = \|\mathcal{I}_h \Pi_H \tilde{m}_H\|_{a,E} \lesssim \sqrt{\beta/\alpha} \|\tilde{m}_H\|_{a,\omega_E}$$

for any $E \in \mathcal{E}_H$, which follows from the definition of $\mathcal{C}_E$ in (4.7) and Lemma 3.4. Finally, using the stability of $\tilde{m}_H$ in (4.9) as well as the limited overlap of the supports ω_E, we obtain

$$\|(\mathcal{R} - \mathcal{R}^\ell)\Pi_H \tilde{m}_H\|_a \lesssim \frac{\max\{1,\beta^{3/2}\}}{\min\{1,\alpha^2\}} \theta^\ell \|f\|_0.$$

We now combine this estimate with (6.3), which yields the assertion. The result is also valid for $\ell \leq 2$, for which only the constant must be suitably adjusted. □

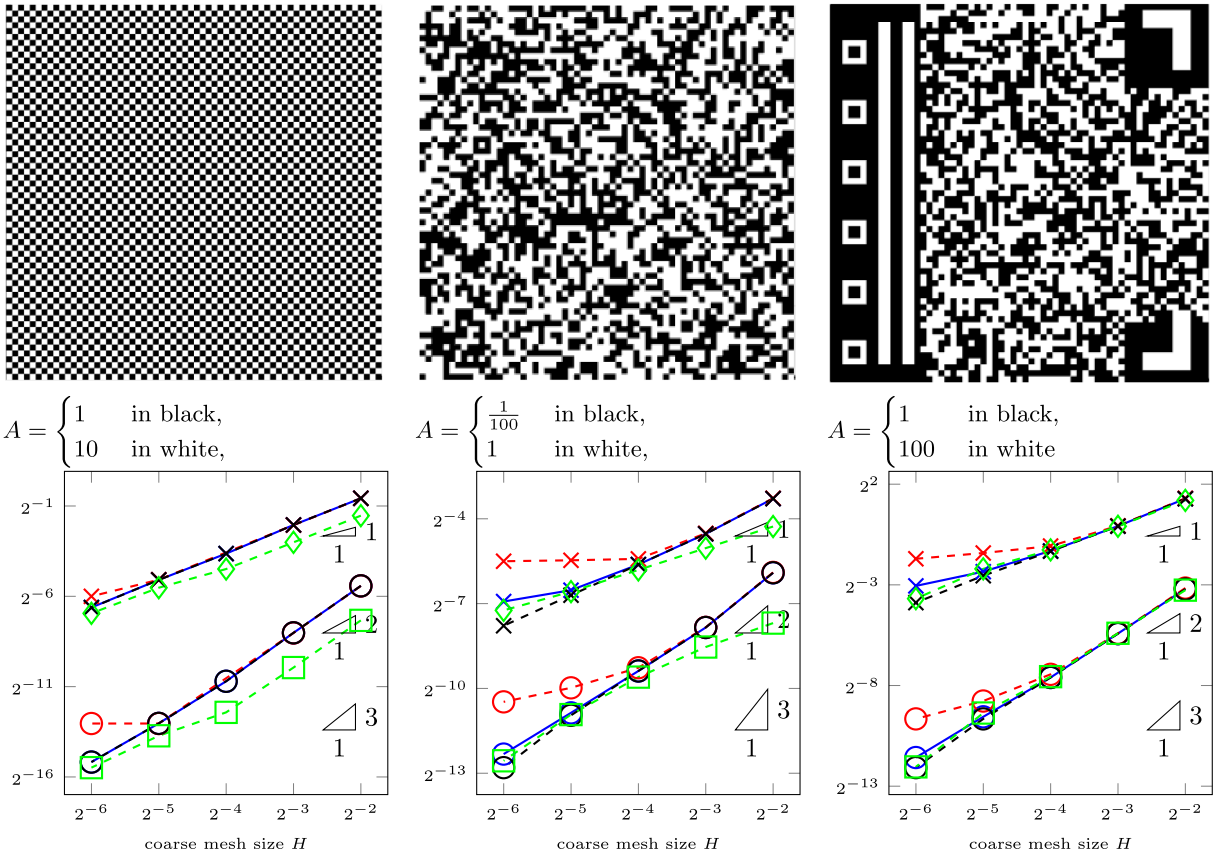

FIGURE 1. Numerical experiments: Checkerboard case in the left column, random case in the middle column, and channel case in the right column. The structure of the diffusion coefficient is illustrated in the top row, the values in black and white cells are given in the second row, and the error plots are presented in the third row. Here, the red dashed lines represent $\ell = 2$, while the blue solid lines correspond to $\ell = 3$, and the black lines refer to $\ell = 4$. Circular marks refer to the L^2 error, and crosses refer to the energy error. The green graphs illustrate simulations with $\ell = 5$, where the contrast has been multiplied by a factor of 10: In the first example, the value of white pixels is replaced by 100, in the second one by 10, and in the third one by 1000. Here, squares refer to the L^2 error, and diamonds to the energy error in this case.

7. NUMERICAL EXPERIMENTS

We investigate the behavior of the practical multiscale method (6.2) to approximate the solution to (2.1) with A as defined in Figure 1 (first and second row). We refer to these as the checkerboard, random, and channel case, respectively. The coefficients are piecewise constant on a domain partition into $2^6 \times 2^6$ small squares, and each cell is either black or white. In the checkerboard case, we set $f(x) = 5\pi^2 \sin(2\pi x_1)\cos(2\pi x_2)$, while $f = 1$ is chosen in the random case, and $f = 100$ in the channel case.

The theoretical results of our work quantify the error of the multiscale solution compared to a very fine reference HDG solution based on a mesh with mesh size h that resolves all the oscillations in the coefficient. For our numerical examples, we set $h = 2^{-8}$ and work with the linear LDG-H method with $\tau = \frac{100}{h}$. In Figure 1 (bottom row), we plot the errors of solutions to (6.2) for different coarse mesh sizes H, $p = 1$, and $\ell = 2, 3, 4$. We present the errors in the energy norm (crosses) and the L^2 norm (circles). We also include the same experiments

with a larger contrast (additional factor 10) and $\ell = 5$ (green graphs, squares depict the L^2 error and diamonds the energy error). From the theory, one would expect worse behavior with higher contrast, which is not observed here. Only the parameter ℓ needs to be slightly adjusted. Generally, we observe that if ℓ is too small, the error curve stagnates at some point if H is decreased, which is in line with the theoretical result in Theorem 6.1 that states that ℓ needs to be suitably increased with decreasing H to keep the convergence rate.

8. Conclusions

In this manuscript, we have derived and analyzed a localized orthogonal decomposition approach for hybrid discontinuous Galerkin formulations (LDG-H, RT-H, or BDM-H) to discretize elliptic PDEs with highly oscillatory rough coefficients. The main obstacle in constructing such a method is the requirement of stable mesh transfer operators to deal with the non-nestedness of the discrete spaces. We have presented an appropriate choice of such operators and proved an optimal first-order error estimate for an idealized globally defined multiscale method. Further, the localization of the method has been analyzed, and its practical behavior has been illustrated in numerical examples.

We believe the presented analysis is only a first step towards the reliable combination of hybrid discontinuous Galerkin methods and the localized orthogonal decomposition strategy. In particular, the aim is to extend the scheme and the corresponding analysis to enable higher-order convergence rates exploiting higher-order ansatz spaces of the HDG postprocessing strategies.

Acknowledgments

A. Rupp has been supported by the Academy of Finland's grant number 350101 *Mathematical models and numerical methods for water management in soils*, grant number 354489 *Uncertainty quantification for PDEs on hypergraphs*, grant number 359633 *Localized orthogonal decomposition for high-order, hybrid finite elements*, Business Finland's project number 539/31/2023 *3D-Cure: 3D printing for personalized medicine and customized drug delivery*, and the Finnish *Flagship of advanced mathematics for sensing, imaging and modeling*, decision number 358944. R. Maier acknowledges support from the German Academic Exchange Service (DAAD), project number 57711336. P. Lu is supported by the National Key R&D Program of China (No. 2024YFA1016300).

References

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

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

[3] G. Allaire and R. Brizzi, A multiscale finite element method for numerical homogenization. *Multiscale Model. Simul.* **4** (2005) 790–812.

[4] R. Altmann, P. Henning and D. Peterseim, Numerical homogenization beyond scale separation. *Acta Numer.* **30** (2021) 1–86.

[5] R. Araya, C. Harder, D. Paredes and F. Valentin, Multiscale hybrid-mixed method. *SIAM J. Numer. Anal.* **51** (2013) 3505–3531.

[6] T. Arbogast, G. Pencheva, M.F. Wheeler and I. Yotov, A multiscale mortar mixed finite element method. *Multiscale Model. Simul.* **6** (2007) 319–346.

[7] D.N. Arnold and F. Brezzi, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates. *ESAIM Math. Model. Numer. Anal.* **19** (1985) 7–32.

[8] I. Babuška and R. Lipton, Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. *Multiscale Model. Simul.* **9** (2011) 373–406.

[9] I. Babuška and J.E. Osborn, Generalized finite element methods: their performance and their relation to mixed methods. *SIAM J. Numer. Anal.* **20** (1983) 510–536.

[10] I. Babuška, G. Caloz and J.E. Osborn, Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.* **31** (1994) 945–981.

[11] G.R. Barrenechea, A.T.A. Gomes and D. Paredes, A multiscale hybrid method. *SIAM J. Sci. Comput.* **46** (2024) A1628–A1657.

[12] D. Boffi, F. Brezzi and M. Fortin, Mixed Finite Element Methods and Applications. Springer, Heidelberg (2013).

[13] S.C. Brenner, Poincaré–Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Anal.* **41** (2003) 306–324.

[14] T. Chaumont-Frelet, A. Ern, S. Lemaire and F. Valentin, Bridging the multiscale hybrid-mixed and multiscale hybrid high-order methods. *ESAIM Math. Model. Numer. Anal.* **56** (2022) 261–285.

[15] H. Chen, P. Lu and X. Xu, A robust multilevel method for hybridizable discontinuous Galerkin method for the Helmholtz equation. *J. Comput. Phys.* **264** (2014) 133–151.

[16] P.G. Ciarlet, Basic error estimates for elliptic problems, in Finite Element Methods (Part 1). Vol. 2 of *Handbook of Numerical Analysis*. Elsevier (1991) 17–351.

[17] M. Cicuttin, A. Ern and S. Lemaire, A hybrid high-order method for highly oscillatory elliptic problems. *Comput. Methods Appl. Math.* **19** (2019) 723–748.

[18] B. Cockburn and J. Gopalakrishnan, Error analysis of variable degree mixed methods for elliptic problems via hybridization. *Math. Comput.* **74** (2005) 1653–1677.

[19] B. Cockburn, J. Gopalakrishnan and R. Lazarov, Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.* **47** (2009) 1319–1365.

[20] B. Cockburn, O. Dubois, J. Gopalakrishnan and S. Tan, Multigrid for an HDG method. *IMA J. Numer. Anal.* **34** (2013) 1386–1425.

[21] D.A. Di Pietro and A. Ern, Mathematical Aspects of Discontinuous Galerkin Methods. *Mathématiques et applications*. Springer, Heidelberg, New York, London (2012).

[22] D.A. Di Pietro, A. Ern and S. Lemaire, An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators. *Comput. Methods Appl. Math.* **14** (2014) 461–472.

[23] Z. Dong, M. Hauck and R. Maier, An improved high-order method for elliptic multiscale problems. *SIAM J. Numer. Anal.* **61** (2023) 1918–1937.

[24] Y. Efendiev, J. Galvis and T.Y. Hou, Generalized multiscale finite element methods (GMsFEM). *J. Comput. Phys.* **251** (2013) 116–135.

[25] Y. Efendiev, R. Lazarov and K. Shi, A multiscale HDG method for second order elliptic equations. Part I. Polynomial and homogenization-based multiscale spaces. *SIAM J. Numer. Anal.* **53** (2015) 342–369.

[26] D. Elfverson, E.H. Georgoulis and A. Målqvist, An adaptive discontinuous Galerkin multiscale method for elliptic problems. *Multiscale Model. Simul.* **11** (2013) 747–765.

[27] D. Elfverson, E.H. Georgoulis, A. Målqvist and D. Peterseim, Convergence of a discontinuous Galerkin multiscale method. *SIAM J. Numer. Anal.* **51** (2013) 3351–3372.

[28] B. Fraeijs de Veubeke, Displacement and equilibrium models in the finite element method, in Stress Analysis. John Wiley & Sons (1965) 145–197.

[29] P. Freese, M. Hauck, T. Keil and D. Peterseim, A super-localized generalized finite element method. *Numer. Math.* **156** (2024) 205–235.

[30] V. Girault and P.A. Raviart, Finite Element Methods for Navier–Stokes Equations. Springer-Verlag, Berlin Heidelberg (1986).

[31] J. Gopalakrishnan, A Schwarz preconditioner for a hybridized mixed method. *Comput. Meth. Appl. Math.* **3** (2003) 116–134.

[32] L. Grasedyck, I. Greff and S. Sauter, The AL basis for the solution of elliptic problems in heterogeneous media. *Multiscale Model. Simul.* **10** (2012) 245–258.

[33] M. Hauck and D. Peterseim, Super-localization of elliptic multiscale problems. *Math. Comput.* **92** (2023) 981–1003.

[34] C. Harder, D. Paredes and F. Valentin. A family of multiscale hybrid-mixed finite element methods for the Darcy equation with rough coefficients. *J. Comput. Phys.* **245** (2013) 107–130.

[35] F. Hellman, P. Henning and A. Målqvist, Multiscale mixed finite elements. *Discrete Contin. Dyn. Syst. Ser. S* **9** (2016) 1269–1298.

[36] P. Henning and D. Peterseim, Oversampling for the multiscale finite element method. *Multiscale Model. Simul.* **11** (2013) 1149–1175.

[37] J.S. Hesthaven, S. Zhang and X. Zhu, High-order multiscale finite element method for elliptic problems. *Multiscale Model. Simul.* **12** (2014) 650–666.

[38] C. Le Bris, F. Legoll and A. Lozinski, MsFEM à la Crouzeix-Raviart for highly oscillatory elliptic problems, in Partial Differential Equations: Theory, Control and Approximation. Springer, Dordrecht (2014) 265–294.

[39] R. Li, P. Ming and F. Tang, An efficient high order heterogeneous multiscale method for elliptic problems. *Multiscale Model. Simul.* **10** (2012) 259–283.
[40] P. Lu, A. Rupp and G. Kanschat, Homogeneous multigrid for HDG. *IMA J. Numer. Anal.* **42** (2022) 3135–3153.
[41] P. Lu, A. Rupp and G. Kanschat, Analysis of injection operators in geometric multigrid solvers for HDG methods. *SIAM J. Numer. Anal.* **60** (2022) 2293–2317.
[42] P. Lu, A. Rupp and G. Kanschat, Homogeneous multigrid for embedded discontinuous Galerkin methods. *BIT Numer. Math.* **62** (2022) 1029–1048.
[43] P. Lu, A. Rupp and G. Kanschat, Two-level Schwarz methods for hybridizable discontinuous Galerkin methods. *J. Sci. Comput.* **95** (2023) 16.
[44] C. Ma and R. Scheichl, Error estimates for discrete generalized FEMs with locally optimal spectral approximations. *Math. Comput.* **91** (2022) 2539–2569.
[45] R. Maier, *Computational multiscale methods in unstructured heterogeneous media*. Ph.D. thesis, University of Augsburg (2020).
[46] R. Maier, A high-order approach to elliptic multiscale problems with general unstructured coefficients. *SIAM J. Numer. Anal.* **59** (2021) 1067–1089.
[47] A. Målqvist and D. Peterseim, Localization of elliptic multiscale problems. *Math. Comput.* **83** (2014) 2583–2603.
[48] A. Målqvist and D. Peterseim, Numerical Homogenization by Localized Orthogonal Decomposition. Vol. 5 of *SIAM Spotlights*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2020).
[49] A.L. Madureira and M. Sarkis, Hybrid localized spectral decomposition for multiscale problems. *SIAM J. Numer. Anal.* **59** (2021) 829–863.
[50] P. Monk, Finite Element Methods for Maxwell's Equations. Oxford University Press (2003).
[51] J. Nečas, Direct Methods in the Theory of Elliptic Equations. Springer Monographs in Mathematics. Springer, Heidelberg (2012). Translated from the 1967 French original by Gerard Tronel and Alois Kufner, Editorial coordination and preface by Šárka Nečasová and a contribution by Christian G. Simader.
[52] H. Owhadi, Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *SIAM Rev.* **59** (2017) 99–149.
[53] H. Owhadi and C. Scovel, Operator-adapted Wavelets, Fast Solvers, and Numerical Homogenization. Vol. 35 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge (2019).
[54] S. Tan, *Iterative solvers for hybridized finite element methods*. Ph.D. thesis, University of Florida (2009).
[55] W. Wang, J. Guzmán and C.-W. Shu, The multiscale discontinuous Galerkin method for solving a class of second order elliptic problems with rough coefficients. *Int. J. Numer. Anal. Model.* **8** (2011) 28–47.
[56] M. Weymuth, *Adaptive local basis for elliptic problems with L^∞-coefficients*. Ph.D. thesis, University of Zurich (2016).

Appendix A. Construction of the lifting operator

A possible operator $S\colon M_h \to V_{\mathrm{disc}}^{\mathrm{c}}$ has been constructed in Section 5.2 of [41] and is motivated by Tan [54], Lemma A.3 of [30] (which guarantees that the operator is well-defined in two spatial dimensions), and Definition 5.46 of [50]. We follow their construction and characterize $S\mu$ in two spatial dimensions *via*

$$\int_e S\lambda\, v \,\mathrm{d}x = \int_e \mathcal{U}\lambda\, v \,\mathrm{d}x \qquad \text{for all } v \in \mathbb{P}_p(e) \text{ and } e \in \mathcal{E}_h, \tag{A.1a}$$

$$\int_F S\lambda\, \eta \,\mathrm{d}\sigma = \int_F \lambda\, \eta \,\mathrm{d}\sigma \qquad \text{for all } \eta \in \mathbb{P}_{p+1}(F) \text{ and } F \in \mathcal{F}_h \tag{A.1b}$$

$$S\lambda(\boldsymbol{x}) = \{\{\lambda(\boldsymbol{x})\}\} \qquad \text{for all vertices } \boldsymbol{x} \text{ in } \mathcal{E}_h, \tag{A.1c}$$

where $\{\{\lambda(\boldsymbol{x})\}\}$ is the mean of all possibly attained values in vertex $\boldsymbol{x}$. The operator S can be extended to three dimensions using (5.11e) of [41] if

$$V_{\mathrm{disc}}^{\mathrm{c}} := \big\{u \in H_0^1(\Omega)\colon u|_E \in \mathbb{P}_{p+d+1}(E) \text{ for all } E \in \mathcal{E}_h\big\}.$$

The following variation of Lemma 5.5 from [41] holds.

Lemma A.1 (Properties of S)**.** *If $\tau h \lesssim 1$ and (LS1)–(LS4) hold, we have that*

$$\begin{aligned}
\mu &= \gamma_h \pi_h S\mu && \text{(trace identity)}\\
\|S\mu\|_{0,E} &\eqsim \|\mu\|_{h,E} && \text{(norm equivalence)},\\
S\gamma_h w &= w && \text{(lifting identity)},\\
|S\mu|_{1,E} &\eqsim \big\|A^{-1}\boldsymbol{\mathcal{Q}}\mu\big\|_{0,E} && \text{(lifting estimate)}.
\end{aligned}$$

for all $\mu \in M_h$ and $w \in \overline{V}_h$.

Proof. The norm equivalence, the lifting identity, and the upper bound of the lifting estimate can be shown completely analogously to Lemma 5.5 of [41]. The trace identity follows immediately from (A.1b). To show the lower bound of the lifting estimate, we observe that for $\mu \in M_h$, $\boldsymbol{v} \in \mathbb{P}_p^d(e)$, and $e \in \mathcal{E}_h$

$$\begin{aligned}
\int_e A^{-1}\boldsymbol{\mathcal{Q}}\mu \cdot \boldsymbol{v} \,\mathrm{d}x &= \int_e \mathcal{U}\mu \,\operatorname{div}\boldsymbol{v} \,\mathrm{d}x - \int_{\partial e} \mu\, v \cdot \boldsymbol{\nu} \,\mathrm{d}\sigma\\
&= \int_e S\mu \,\operatorname{div}\boldsymbol{v} \,\mathrm{d}x - \int_{\partial e} S\mu\, \boldsymbol{v}\cdot\boldsymbol{\nu} \,\mathrm{d}\sigma = -\int_e \nabla S\mu \cdot \boldsymbol{v} \,\mathrm{d}x,
\end{aligned}$$

where the first equality is (3.1a), the second equality is (A.1a) and (A.1b), and the third equality is integration by parts. The above identity implies that $-A^{-1}\boldsymbol{\mathcal{Q}}\mu$ is the L^2-orthogonal projection of $\nabla S\mu$ onto $\mathbb{P}_p^d(e)$, which in turn implies the lower bound of the lifting estimate. □

Appendix B. Poincaré–Friedrichs inequalities for DG and HDG

B.1. Continuous Poincaré–Friedrichs inequality

We first state a continuous version of the Poincaré–Friedrichs inequality generalized to certain non-convex domains.

Theorem B.1. *Consider a domain $\Omega \subset \mathbb{R}^d$ that can be decomposed into $n \in \mathbb{N}$ (not necessarily disjoint) convex domains $\Omega_i \neq \emptyset$ $(i = 1, \dots, n)$. That is, we need to have that $\Omega = \operatorname{int}\bigcup_{i=1}^n \bar{\Omega}_i$. Assume that each Ω_i can be inscribed into a ball around some point x_i with diameter r. Moreover, let $\Gamma_i \subset \partial\Omega_i$ be such that $\mu(\Omega_i)/\mu(\Gamma_i) \leq \delta r$ for some $\delta > 0$ and all $i = 1, \dots, n$. Then for all $u \in H^1(\Omega)$, we have*

$$\|u\|_{L^2(\Omega)}^2 \lesssim r^2\|\nabla u\|_{L^2(\Omega)}^2 + \delta r\|u\|_{L^2(\Gamma)}^2,$$

where $\Gamma = \bigcup_{i=1}^n \Gamma_i$.

Proof. For a convex set, the proof follows the same lines as, *e.g.*, Theorem 1.2 of [51] and can be directly extended to a union of such domains. □

Note that the result covers the L-shaped domain shown in Figure B.1 (left), with a possible decomposition indicated by the dashed line and Γ is highlighted by the two bold lines. The (non-Lipschitz) domain with a crack in Figure B.1 (right) is also covered if, for example, Γ is identical to the crack. In this case, the dashed line indicates a possible decomposition.

FIGURE B.1. Illustration of a non-convex Lipschitz domain (*left*) and a non-convex square with crack (*right*) with possible decomposition.

B.2. Poincaré–Friedrichs inequality for broken Sobolev spaces

As above, let us consider a polygonally bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, which is discretized by a geometrically conforming family of triangulations $(\mathcal{E}_h)_h$, *i.e.*, partitions of Ω into simplices without hanging nodes. We require that any face $F \subset \partial\Omega$ either satisfies $F \subset \Gamma$ or $\mu(F \cap \Gamma) = 0$. The family $(\mathcal{E}_h)_h$ is supposed to be shape-regular, which prevents simplices from deteriorating. That is, all angles are bounded away from zero. For a mesh $\mathcal{E}_h = \{E_1, \dots, E_N\}$ consisting of N elements and an element-wise defined function u_h (smooth enough to have traces), we write $\{\{u_h\}\}$ for the average with respect to a face, and $[\![u_h]\!]$ for the jump, *i.e.*, for $F \subset \partial E_i \cap \partial E_j$ with $i < j$, we have

$$\{\{u_h\}\} = \tfrac{1}{2}(u_{h,i} + u_{h,j}), \qquad [\![u_h]\!] = u_{h,i} - u_{h,j},$$

where $u_{h,k} := u_h|_{T_k}$, $k = 1, \dots N$. Further, for $F \subset \partial\Omega$, we have

$$\{\{u_h\}\} = u_h \quad \text{and} \quad [\![u_h]\!] = 0.$$

Next, we define the broken Sobolev space

$$H^1(\mathcal{E}_h) = \big\{u_h \in L^2(\Omega) : \ u_h|_E \in H^1(E) \ \text{ for all } E \in \mathcal{E}_h\big\}$$

with induced semi-norm and dG jump semi-norm

$$|u_h|^2_{H^1(\mathcal{E}_h)} = \sum_{E\in\mathcal{E}_h} |u_h|^2_{H^1(E)} \quad \text{and} \quad |u_h|^2_{[\![\cdot]\!]} = \sum_{F\in\mathcal{F}_h} \frac{1}{h_F}\int_F [\![u_h]\!]^2 \,\mathrm{d}\sigma,$$

respectively. Once again, $\mathcal{F}_h$ denotes the set of faces of $\mathcal{E}_h$, $h_F = \operatorname{diam}(F)$, $h_E = \operatorname{diam}(E)$ and $h = \max_{E\in\mathcal{E}_h} h_E$. Notably, the dG norm

$$\|u_h\|^2_{\mathrm{dG}} = |u_h|^2_{H^1(\mathcal{E}_h)} + |u_h|^2_{[\![\cdot]\!]}$$

is a genuine norm on $H^1(\mathcal{E}_h)$. It can be understood as the energy norm of many dG schemes or analogous to the H^1-semi-norm for $H^1_0(\Omega)$ functions. We formulate a Poincaré–Friedrichs inequality for this norm.

Theorem B.2. *Let the assumptions of Theorem B.1 hold. Further, let $(\mathcal{E}_h)_h$ be a shape-regular family of geometrically conforming triangulations of Ω into simplices. Then we have for all $u_h \in H^1(\mathcal{E}_h)$ that*

$$\|u_h\|_{L^2(\Omega)} \lesssim \Big(h + \sqrt{\delta r h}\Big)\|u_h\|_{\mathrm{dG}} + r|u_h|_{H^1(\mathcal{E}_h)} + \sqrt{\delta r}\|u_h\|_{L^2(\Gamma)}. \tag{B.1}$$

Proof. The proof mainly follows from the arguments in [13]. In particular, the left-hand side of (B.1) is split into multiple contributions using the triangle inequality. These different terms are individually bounded, and Theorem B.1 is employed. □

B.3. Poincaré–Friedrichs inequality for broken skeleton spaces

Recall that the skeleton space $M_h \subset L^2(\mathcal{F}_h)$ is given as an abstract space with norm

$$\|m\|^2_h = \sum_{E\in\mathcal{E}_h} \frac{\mu(E)}{\mu(\partial E)}\int_{\partial E} m^2 \,\mathrm{d}\sigma.$$

We emphasize that this norm can be directly applied to broken Sobolev functions as well, *i.e.*,

$$\|u_h\|_h^2 = \sum_{E\in\mathcal{E}_h} \frac{\mu(E)}{\mu(\partial E)} \int_{\partial E} u_h|_T^2 \,\mathrm{d}\sigma \qquad \text{for all } u_h \in H^1(\mathcal{E}_h).$$

If $u_h \in \mathbb{P}_p(\mathcal{E}_h) = \{v_h \in H^1(\mathcal{E}_h) : v_h|_E \in \mathbb{P}_p(E) \text{ for all } E \in \mathcal{E}_h\}$, we even have that

$$\|u_h\|_h \lesssim \|u_h\|_{L^2(\Omega)} \tag{B.2}$$

as a direct consequence of Lemma 1.46 from [21]. For the following result, we work with a generalized concept of the local solvers as described in Section 3. We only assume that the (non-surjective) local solvers

$$(\mathcal{U}, \boldsymbol{\mathcal{Q}})\colon M_h \ni m \mapsto (\mathcal{U}m, \boldsymbol{\mathcal{Q}}m) \in H^1(\mathcal{E}_h) \times L^2(\Omega)^d$$

satisfy (LS3) and the following slightly adapted version of (LS1), *i.e.*

$$\|\mathcal{U}m - m\|_h^2 \lesssim \sum_{E\in\mathcal{E}_h} h_E^2 \big\|A^{-1}\boldsymbol{\mathcal{Q}}m\big\|_{L^2(E)}^2, \tag{LS1*}$$

where h_E is the diameter of E. Importantly, (LS3) and a variant of (LS1*) hold for the LDG-H, RT-H, and BDM-H methods according to Section 6 of [41]. Our condition (LS1*) can be proved by using their arguments in concert with Lemmas 3.1 & 3.4 of [20]. The following result now quantifies a Poincaré–Friedrichs inequality for HDG methods.

Theorem B.3 (broken Poincaré–Friedrichs inequality)**.** *Suppose that the assumptions of Theorem B.2 hold. If $\mathcal{U}$ and $\boldsymbol{\mathcal{Q}}$ satisfy* (LS1*) *and* (LS3)*, we have*

$$\|m\|_h + \|\mathcal{U}m\|_{L^2(\Omega)} \lesssim \Big(r + h + \sqrt{\delta r h}\Big)\big\|A^{-1}\boldsymbol{\mathcal{Q}}m\big\|_{L^2(\Omega)} + \sqrt{\delta r}\|m\|_{L^2(\Gamma)}$$

for all $m \in M_h$.

Proof. The estimates (LS1*) and (B.2) yield

$$\|m\|_h \le \|m - \mathcal{U}m\|_h + \|\mathcal{U}m\|_h \lesssim h\big\|A^{-1}\boldsymbol{\mathcal{Q}}m\big\|_{L^2(\Omega)} + \|\mathcal{U}m\|_{L^2(\Omega)}.$$

Next, we apply Theorem B.2, and afterwards (LS3) and (LS1*) to obtain

$$\begin{aligned}
\|\mathcal{U}m\|_{L^2(\Omega)}^2 &\lesssim \big(h^2 + \delta r h\big)\|\mathcal{U}m\|_{\mathrm{dG}}^2 + r^2|\mathcal{U}m|_{H^1(\mathcal{E}_h)}^2 + \delta r\|\mathcal{U}m\|_{L^2(\Gamma)}^2 \\
&\lesssim \big(h^2 + r^2 + \delta r h\big)|\mathcal{U}m|_{H^1(\mathcal{E}_h)}^2 + \delta r\|\mathcal{U}m\|_{L^2(\Gamma)}^2 + (h+\delta r)\sum_{F\in\mathcal{F}_h}\int_F [\![\mathcal{U}m]\!]^2\,\mathrm{d}\sigma \\
&\lesssim \big(h^2 + r^2 + \delta r h\big)\big\|A^{-1}\boldsymbol{\mathcal{Q}}m\big\|_{L^2(\mathcal{E}_h)}^2 + \delta r \min\Big\{\|m\|_{L^2(\Gamma)}^2, \|\mathcal{U}m\|_{L^2(\Gamma)}^2\Big\} \\
&\quad + (h+\delta r)\sum_{E\in\mathcal{E}_h}\int_{\partial E}(\mathcal{U}m - m)^2\,\mathrm{d}\sigma.
\end{aligned}$$

The last inequality exploits the triangle inequality, and adding more positive integrals only enlarges the right-hand side. Note that we obtain the minimum in the last step as one can keep the term $\delta r\|\mathcal{U}m\|_{L^2(\Gamma)}^2$ or add $\pm m$. In the latter case, the difference term enters in the last term. The result then follows from the definition of $\|\mathcal{U}m - m\|_h$ and (LS1*). □

For HDG bilinear forms, which have the general form

$$a(m,\mu) = \int_\Omega A^{-1}\boldsymbol{\mathcal{Q}}m \cdot \boldsymbol{\mathcal{Q}}\mu\,\mathrm{d}x + s(m,\mu)$$

with a symmetric positive semi-definite stabilizer/penalty term s, the above results clearly show that also

$$\|m\|_h^2 + \|\mathcal{U}m\|_{L^2(\Omega)}^2 \lesssim \alpha^{-1}\Big(r + h + \sqrt{\delta r h}\Big)^2 a(m,m) + \delta r\|m\|_{L^2(\Gamma)}^2$$

for all $m \in M_h$. Further, the term $\|m\|_{L^2(\Gamma)}$ vanishes if $m = 0$ on $\Gamma \subset \partial\Omega$.

Remark B.4 (Covered methods)**.** The results in this section are guaranteed to hold for the LDG-H, RT-H, and BDM-H methods. If (LS1*) and (LS3) hold for HHO, the results of this section readily apply to HHO as well.