Numer. Math. (2015) 131:643–682
DOI 10.1007/s00211-015-0703-z

Numerische
Mathematik

On multiscale methods in Petrov–Galerkin formulation

Daniel Elfverson · Victor Ginting ·
Patrick Henning

Received: 22 May 2014 / Revised: 17 November 2014 / Published online: 11 January 2015

Abstract In this work we investigate the advantages of multiscale methods in Petrov–Galerkin (PG) formulation in a general framework. The framework is based on a localized orthogonal decomposition of a high dimensional solution space into a low dimensional multiscale space with good approximation properties and a high dimensional remainder space, which only contains negligible fine scale information. The multiscale space can then be used to obtain accurate Galerkin approximations. As a model problem we consider the Poisson equation. We prove that a Petrov–Galerkin formulation does not suffer from a significant loss of accuracy, and still preserve the convergence order of the original multiscale method. We also prove inf-sup stability of a PG continuous and a discontinuous Galerkin finite element multiscale method. Furthermore, we demonstrate that the Petrov–Galerkin method can decrease the com-

D. Elfverson and P. Henning were partially supported by the Göran Gustafsson Foundation and the Swedish Research Council.

D. Elfverson
Department of Information Technology, Uppsala University,
Box 337, 751 05 Uppsala, Sweden
e-mail: daniel.elfverson@it.uu.se

V. Ginting
Department of Mathematics, University of Wyoming, Laramie, WY 82071, USA
e-mail: vginting@uwyo.edu

P. Henning
Section de Mathématiques, École polytechnique fédérale de Lausanne,
1015 Lausanne, Switzerland

Present Address:
P. Henning (✉)
University of Münster, 48149 Münster, Germany
e-mail: patrick.henning@uni-muenster.de

putational complexity significantly, allowing for more efficient solution algorithms. As another application of the framework, we show how the Petrov–Galerkin framework can be used to construct a locally mass conservative solver for two-phase flow simulation that employs the Buckley–Leverett equation. To achieve this, we couple a PG discontinuous Galerkin finite element method with an upwind scheme for a hyperbolic conservation law.

Mathematics Subject Classification 35J15 · 65N12 · 65N30 · 76S05

1 Introduction

In this contribution we consider linear elliptic problems with a heterogenous and highly variable diffusion coefficient A as arisen often in hydrology or in material sciences. In the following, we are looking for u which solves

$$\begin{aligned} -\nabla \cdot A\nabla u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

in a weak sense. Here, we denote

(A1) $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, a bounded Lipschitz domain with a piecewise polygonal boundary,

(A2) $f \in L^2(\Omega)$ a source term, and,

(A3) $A \in L^\infty(\Omega, \mathbb{R}^{d\times d}_{sym})$ a symmetric matrix-valued function with uniform spectral bounds $\beta_0 \geq \alpha_0 > 0$, $\sigma(A(x)) \subset [\alpha_0, \beta_0]$ for almost all $x \in \Omega$. We call the ratio β_0/α_0 the *contrast* of A.

Under assumptions (A1)–(A3) and by the Lax–Milgram theorem, there exists a unique weak solution $u \in H^1_0(\Omega)$ to

$$a(u, v) = (f, v) \quad \text{for all } v \in H^1_0(\Omega), \tag{1}$$

where

$$a(v, w) := \int_\Omega A\nabla v \cdot \nabla w \quad \text{and} \quad (v, w) := (v, w)_{L^2(\Omega)}.$$

The problematic term in the equation is the diffusion matrix A, which is known to exhibit very fast variations on a very fine scale (i.e. it has a multiscale character). These variations can be highly heterogenous and unstructured, which is why it is often necessary to resolve them globally by an underlying computational grid that matches the said heterogeneity. Using standard finite element methods, this results in high dimensional solution spaces and hence an enormous computational demand, which often cannot be handled even by today's computing technology. Consequently, there is a need for alternative methods, so called multiscale methods, which can either operate below linear computational complexity by using local representative elements (cf. [1,2,17,18,23,36,40]) or which can split the original problem into very localized subproblems that cover Ω but that can be solved cheaply and independent from each other (cf. [5,8,11,12,16,25,27,28,31,33,37,38]).

In this paper, we focus on a rather recent approach called localized orthogonal decomposition (LOD) that was introduced by Målqvist and Peterseim [35] and further generalized in [19,24].

We consider a coarse space V_H, which is low-dimensional but possibly inadequate for finding a reliable Galerkin approximation to the multiscale solution of problem (1). The idea of the method is to start from this coarse space and to update the corresponding set of basis functions step-by-step to improve the approximation properties of the space. In a summarized form, this can be described in four steps: (1) define a (quasi) interpolation operator I_H from $H^1_0(\Omega)$ onto V_H, (2) information in the kernel of the interpolation operator is considered to be negligible (having a small L^2-norm), (3) hence define the space of negligible information by the kernel of this interpolation, i.e. $W := \text{kern}(I_H)$, and (4) find the orthogonal complement of W with respect to a scalar product $a_h(\cdot,\cdot)$, where $a_h(\cdot,\cdot)$ describes a discretization of the problem to solve. In many cases, it can be shown, that this (low dimensional) orthogonal complement space has very accurate approximation properties with respect to the exact solution. Typically, the computation of the orthogonal decomposition is localized to small patches in order to reduce the computational complexity.

So far, the concept of the LOD has been successfully applied to nonlinear elliptic problems [20], eigenvalue problems [34] and the nonlinear Schrödinger equation [21]. Furthermore, it was combined with a discontinuous Galerkin method [13,14] and extended to the setting of partition of unity methods [22].

In this work, we are concerned with analyzing the LOD framework in Petrov–Galerkin formulation, i.e. for the case that the discrete trial and test spaces are not identical. We show that an LOD method in Petrov–Galerkin formulations still preserves the convergence rates of the original formulation of the method. At the same time, the new method can exhibit significant advantages, such as decreased computational complexity and mass conservation properties. In this paper, we discuss these advantages in detail; we give examples for realizations and present numerical experiments. In particular, we apply the proposed framework to design a locally conservative multiscale solver for the simulation of two-phase flow models as governed by the Buckley–Leverett equation. We remark that employing Petrov–Galerkin variational frameworks in the construction and analysis of multiscale methods for solving elliptic problems in heterogeneous media has been investigated in the past, see for example [16,26].

The rest of the paper is organized as follows. Section 2 lays out the setting and notation for the formulation of the multiscale methods that includes the description of two-grid discretization and the LOD. In Sect. 3, we present the multiscale methods based on the LOD framework, starting from the usual Galerkin variational equation and concentrating further on the Petrov–Galerkin variational equation that is the main contribution of the paper. We establish in this section that the Petrov–Galerkin LOD (PG-LOD) exhibits the same convergence behavior as the usual Galerkin LOD (G-

LOD). Furthermore, we draw a contrast in the aspect of practical implementation that makes up a strong advantage of PG-LOD in relative comparison to G-LOD. The other advantage of the PG-LOD which cannot be achieved with G-LOD is the ability to produce a locally conservative flux field at the elemental level when discontinuous finite element is utilized. We also discuss in this section an application of the PG-LOD for solving the pressure equation in the simulation of two-phase flow models to demonstrate this particular advantage. Section 4 gives two sets of numerical experiment: one that confirms the theoretical finding and the other demonstrating the application of PG-LOD in the two-phase flow simulation. We present the proofs of the theoretical findings in Sect. 5.

2 Discretization

In this section we introduce notations that are required for the formulation of the multiscale methods.

2.1 Abstract two-grid discretization

We define two different meshes on Ω. The first mesh is a ‘coarse mesh’ and is denoted by $\mathcal{T}_H$, where $H > 0$ denote the maximum diameter of all elements of $\mathcal{T}_H$. The second mesh is a ‘fine mesh’ denoted by $\mathcal{T}_h$ with h representing the maximum diameter of all elements of $\mathcal{T}_h$. By ‘fine’ we mean that any variation of the coefficient A is resolved within this grid, leading to a high dimensional discrete space that is associated with this mesh. The mesh $\mathcal{T}_h$ is assumed to be a (possibly non-uniform) refinement of $\mathcal{T}_H$. Furthermore, both grids are shape-regular and conforming partitions of Ω and we assume that $h < H/2$. For the subsequent methods to make sense, we also assume that each element of $\mathcal{T}_H$ is at least twice uniformly refined to create $\mathcal{T}_h$. The set of all Lagrange points (vertices) of $\mathcal{T}_\star$ is denoted by $\mathcal{N}_\star$, and the set of interior Lagrange points is denoted by $\mathcal{N}_\star^0$, where $\star$ is either H or h.

Now we consider an abstract discretization of the exact problem (1). For this purpose, we let V_h denote a high dimensional discrete space in which we seek an approximation u_h of u. A simple example would be the classical $P1$ Lagrange finite element space associated with $\mathcal{T}_h$. However, note that we do not assume that V_h is a subspace of $H_0^1(\Omega)$. In fact, later we give an example for which V_h consists of non-continuous piecewise linear functions. Next, we assume that we are interested in solving a fine scale problem, that can be characterized by a scalar product $a_h(\cdot,\cdot)$ on V_h. Accordingly, a method on the coarse scale can be described by some $a_H(\cdot,\cdot)$, which we specify by assuming

(A4) $a_\star(\cdot,\cdot)$ is a scalar product on $V_\star$ where $\star$ is either h or H.

This allows us to define the abstract reference problem stated below.

Definition 1 (*Fine scale reference problem*) We call $u_h \in V_h$ the fine scale reference solution if it solves

$$a_h(u_h, v_h) = (f, v_h)_{L^2(\Omega)} \quad \text{for all } v_h \in V_h, \tag{2}$$

where $a_h(\cdot,\cdot)$ ‘describes the method’. It is implicitly assumed that problem (2) is of tremendous computational complexity and cannot be solved by available computing resources i n a convenient time.

A simple example of $a_h(\cdot,\cdot)$ is $a_h(v_h, w_h) = a_H(v_h, w_h) = a(v_h, w_h)$. A more complex example is the $a_h(\cdot,\cdot)$ that stems from a discontinuous Galerkin approximation, in which case $a_h(\cdot,\cdot)$ is different from $a_H(\cdot,\cdot)$. The goal is to approximate problem (2) by a new problem that reaches a comparable accuracy but one that can be solved with a significantly lower computational demand.

2.2 Localized orthogonal decomposition

In this subsection, we introduce the notation that is required in the formulation of the multiscale method. In particular, we introduce an orthogonal decomposition of the high dimensional solution space V_h into the orthogonal direct sum of a low dimensional space with good approximation properties and a high dimensional remainder space. For this purpose, we make the following abstract assumptions.

(A5) $|||\cdot|||_h$ denotes a norm on V_h that is equivalent to the norm that is induced by $a_h(\cdot,\cdot)$, hence there exist generic constants $0 < \alpha \leq \beta$ such that

$$\alpha|||v_h|||_h^2 \leq a_h(v_h, v_h) \quad \text{and} \quad a_h(v_h, w_h) \leq \beta|||v_h|||_h|||w_h|||_h$$

for all $v_h, w_h \in V_h$. In the same way, $|||\cdot|||_H$ denotes a norm on V_H (equivalent to the norm induced by $a_H(\cdot,\cdot)$). Furthermore, we let $C_{H,h}$ denote the constant with $|||v|||_H \leq C_{H,h}|||v|||_h$ for all $v \in V_h$. Note that $C_{H,h}$ might degenerate for $h \to 0$.

(A6) The coarse space $V_H \subset V_h$ is a low dimensional subspace of V_h that is associated with $\mathcal{T}_H$.

(A7) Let $I_H : V_h \to V_H$ be an L^2-stable quasi-interpolation (or projection) operator with the properties

- there exists a generic constant C_{I_H} (only depending on the shape regularity of $\mathcal{T}_H$ and $\mathcal{T}_h$) such that for all $v_h \in V_h$ and $v_H \in V_H$ it holds $\|v_h - I_H(v_h)\|_{L^2(\Omega)} \leq C_{I_H} H|||v_h|||_h$; $|||I_H(v_h)|||_H \leq C_{I_H}|||v_h|||_h$; $\|v_H - I_H(v_H)\|_{L^2(\Omega)} \leq C_{I_H} H|||v_H|||_H$ and $\|I_H(v_H)\|_{L^2(\Omega)} \leq C_{I_H}|||v_H|||_H$,
- the restriction $(I_H)_{|V_H}$ is an isomorphism with $|||\cdot|||_H$-stable inverse, i.e. we have $v_H = (I_H \circ (I_H|_{V_H})^{-1})(v_H)$ for $v_H \in V_H$ and there exists a generic $C_{I_H^{-1}}$ such that for all $v_H \in V_H$ it holds $|||(I_H|_{V_H})^{-1}(v_H)|||_H \leq C_{I_H^{-1}}|||v_H|||_H$.

Typically, L^2-projections onto V_H can be verified to fulfill assumption (A7). Similarly, I_H can be a quasi-interpolation of the Clément-type that is related to the L^2-projection. An example for this case is given in Eq. (13) below. Alternatively, I_H can be also constructed from local L^2-projections as it is done for the classical Clément interpolation. Nodal interpolations typically do not satisfy (A7).

Using the assumption that $(I_H)_{|V_H} : V_H \rightarrow V_H$ is an isomorphism (i.e. assumption (A7)), a splitting of the space V_h is given by the direct sum

$$V_h = V_H \oplus W_h, \quad \text{with } W_h := \{v_h \in V_h | I_H(v_h) = 0\}. \tag{3}$$

Observe that the 'remainder space' W_h contains all fine scale features of V_h that cannot be expressed in the coarse space V_H.

Next, consider the $a_h(\cdot, \cdot)$-orthogonal projection $P_h : V_h \rightarrow W_h$ that fulfills:

$$a_h(P_h(v_h), w_h) = a_h(v_h, w_h) \quad \text{for all } w_h \in W_h. \tag{4}$$

Since $V_h = V_H \oplus W_h$, we have that $V_\Omega^{\text{ms}} := \text{kern}(P_h) = (1 - P_h)(V_H)$ induces the $a_h(\cdot, \cdot)$-orthogonal splitting

$$V_h = V_\Omega^{\text{ms}} \oplus W_h.$$

Note that V_Ω^{ms} is a low dimensional space in the sense that it has the same dimension as V_H. As shown for several applications (cf. [20,21,34]) the space V_Ω^{ms} has very rich approximation properties in the $|||\cdot|||_h$-norm. However, it is very expensive to assemble V_Ω^{ms}, which is why it is practically necessary to localize the space W_h (respectively localize the projection). This is done using admissible patches of the following type.

Definition 2 (*Admissible patch*) For any coarse element $T \in \mathcal{T}_H$, we say that the open and connected set $U(T)$ is an *admissible patch* of T, if $T \subset U(T) \subset \Omega$ and if it consists of elements from the fine grid, i.e.

$$U(T) = \text{int} \bigcup_{\tau \in \mathcal{T}_h^U} \overline{\tau}, \quad \text{where } \mathcal{T}_h^U \subset \mathcal{T}_h.$$

It is now relevant to define the restriction of W_h to an admissible patch $U(T) \subset \Omega$ by

$$\mathring{W}_h(U(T)) := \{v_h \in W_h | v_h = 0 \quad \text{in } \Omega \backslash U(T)\}.$$

A general localization strategy for the space V_Ω^{ms} can be described as follows (see [19] for a special case of this localization and [35] for a different localization strategy).

Definition 3 (*Localization of the solution space*) Let the bilinear form $a_h^T(\cdot, \cdot)$ be a localization of $a_h(\cdot, \cdot)$ on $T \in \mathcal{T}_H$ in the sense that

$$a_h(v_h, w_h) = \sum_{T \in \mathcal{T}_H} a_h^T(v_h, w_h), \tag{5}$$

where $a_h^T(\cdot,\cdot)$ acts only on T or a small environment of T. Let furthermore $U(T)$ be an admissible patch associated with $T \in \mathcal{T}_H$. Let $Q_h^T : V_h \to \mathring{W}_h(U(T))$ be a local correction operator that is defined as finding $Q_h^T(\phi_h) \in \mathring{W}_h(U(T))$ satisfying

$$a_h(Q_h^T(\phi_h), w_h) = -a_h^T(\phi_h, w_h) \quad \text{for all } w_h \in \mathring{W}_h(U(T)), \tag{6}$$

where $\phi_h \in V_h$. The global corrector is given by

$$Q_h(\phi_h) := \sum_{T \in \mathcal{T}_H} Q_h^T(\phi_h). \tag{7}$$

A (localized) generalized finite element space is defined as

$$V^{\mathrm{ms}} := \{\Phi_H + Q_h(\Phi_H) | \, \Phi_H \in V_H\}.$$

The variational formulation (6) is called the corrector problem associated with $T \in \mathcal{T}_H$. Solvability of each of these problems is guaranteed by the Lax–Milgram theorem. By its nature, the system matrix corresponding to (6) is localized to the patch $U(T)$ since the support of w_h is in $U(T)$. Furthermore, each of (6) pertaining to $T \in \mathcal{T}_H$ is designed to be elementally independent and thus attributing to its immediate parallelizability. The corrector problems are solved in a preprocessing step and can be reused for different source terms and for different realization of the LOD methods. Since V^{ms} is a low dimensional space with locally supported basis functions, solving a problem in V^{ms} is rather inexpensive. Normally, the solutions $Q_h^T(\phi_h)$ of (6) decays exponentially to zero outside of T. This is the reason why we can hope for good approximations even for small patches $U(T)$. Later, we quantify this decay by an abstract assumption (which is known to hold true for many relevant applications).

Remark 1 If $U(T) = \Omega$ for all $T \in \mathcal{T}_H$, then $Q_h = -P_h$, where P_h is the orthogonal projection given by (4). In this sense, V^{ms} is localization of the space V_Ω^{ms}. This can be verified using (5), which yields for all $w_h \in W_h$

$$a_h(\phi_h + Q_h(\phi_h), w_h) = \sum_{T \in \mathcal{T}_H} \Big(a_h^T(\phi_h, w_h) + a_h(Q_h^T(\phi_h), w_h) \Big) = 0.$$

By uniqueness of the projection, we conclude $Q_h = -P_h$.

The above setting is used to construct the multiscale methods utilizing the LOD method as e.g. done in [19,35] for the standard finite element formulation and a corresponding Petrov–Galerkin formulation.

3 Methods and properties

In this section, we state the LOD in Galerkin and in Petrov–Galerkin formulation along with their respective a priori error estimates and the inf-sup stability. In the last part of this section, we give two explicit examples and discuss the advantages of the Petrov–Galerkin formulation. Subsequently we use the notation $a \lesssim b$ to abbreviate $a \leq Cb$, where C is a constant that is independent of the mesh sizes H and h; and which is independent of the possibly rapid oscillations in A.

In order to state proper a priori error estimates, we describe the notion of 'patch size' and how the size of $U(T)$ affects the final approximation. All the stated theorems on the error estimates of the LOD methods are proved in Sect. 5.

Definition 4 (*Patch size*) Let $k \in \mathbb{N}_{>0}$ be fixed. We define patches $U(T)$ that consist of the element T and k-layers of coarse element around it. For all $T \in \mathcal{T}_H$, we define element patches in the coarse mesh $\mathcal{T}_H$ by

$$\begin{aligned} U_0(T) &:= T, \\ U_k(T) &:= \cup \left\{ T' \in \mathcal{T}_H | T' \cap U_{k-1}(T) \neq \emptyset \right\} \quad k = 1, 2, \ldots. \end{aligned} \tag{8}$$

The above concept of patch sizes and patch shapes can be also generalized. See for instance [22] for a LOD that is purely based on partitions of unity. Using Definition 4, we make an abstract assumption on the decay of the local correctors $Q_h^T(\Phi_H)$ for $\Phi_H \in V_H$:

(A8) Let $Q_h^{\Omega,T}(\Phi_H)$ be the *optimal* local corrector using $U(T) = \Omega$ that is defined according to (6) and let $Q_h^{\Omega}(\Phi_H) := \sum_{T \in \mathcal{T}_H} Q_h^{\Omega,T}(\Phi_H)$. Let $k \in \mathbb{N}_{>0}$ and for all $T \in \mathcal{T}_H$ let $U(T) = U_k(T)$ as in Definition 4. Then there exists $p \in \{0, 1\}$ and a generic constant $0 < \theta < 1$ that can depend on the contrast, but not on H, h or the variations of A such that for all $\Phi_H \in V_H$,

$$|||(Q_h - Q_h^{\Omega})(\Phi_H)|||_h^2 \lesssim k^d \theta^{2k} (1/H)^{2p} \, |||\Phi_H + Q_h^{\Omega}(\Phi_H)|||_h^2, \tag{9}$$

where $Q_h(\Phi_H)$ is given by (7) for $U(T) = U_k(T)$.

Assumption (A8) quantifies the decay of local correctors, by stating that the solutions of the local corrector problems decay exponentially to zero outside of T. This is central for all a priori error estimates. For continuous Galerkin methods, we can obtain the optimal order $p = 0$ for the exponent in (9). This means, that the $(1/H)$-term fully vanishes. However, depending on the localization strategy [i.e. how $Q_h(\Phi_H)$ is computed] it is also possible that p takes the value 1 and that hence a pollution term of order $(1/H)$ arises (see [19, Remark 3.8] for a discussion). For discontinuous Galerkin methods, the optimal known order is $p = 1$. However, even for this case it is known that the $(1/H)$-term is rapidly overtaken by the decay, leading purely to slightly larger patch sizes (see e.g. [35]).

3.1 Galerkin LOD

This method was originally proposed in [35]: find $u_H^{\text{G-LOD}} \in V^{\text{ms}}$ that satisfies

$$a_h \left(u_H^{\text{G-LOD}}, \Phi^{\text{ms}} \right) = (f, \Phi^{\text{ms}}) \quad \text{for all } \Phi^{\text{ms}} \in V^{\text{ms}}. \tag{10}$$

Theorem 1 (A priori error estimate for Galerkin LOD) *Assume (A1)–(A8). Given a positive $k \in \mathbb{N}_{>0}$, let for all $T \in \mathcal{T}_H$ the patch $U(T) = U_k(T)$ be defined as in (8) and let $u_H^{\text{G-LOD}} \in V^{\text{ms}}$ be as governed by (10). Let $u_h \in V_h$ be the fine scale reference solution governed by (2). Then, the following a priori error estimate holds true*

$$\begin{aligned}&\left\| u_h - ((I_H|_{V_H})^{-1} \circ I_H)(u_H^{\textit{G-LOD}}) \right\|_{L^2(\Omega)} + \left|\left|\left| u_h - u_H^{\textit{G-LOD}} \right|\right|\right|_h \\ &\quad \lesssim (H + (1/H)^p k^{d/2} \theta^k) \| f \|_{L^2(\Omega)},\end{aligned} \tag{11}$$

where $0 < \theta < 1$ and $p \in \{0, 1\}$ are the generic constants in (A8).

The term $((I_H|_{V_H})^{-1} \circ I_H)(u_H^{\text{G-LOD}})$ describes the coarse part (resulting from V_H) of $u_H^{\text{G-LOD}}$ and thus is numerically homogenized (the oscillations are averaged out). In this sense, we can say that $u_H^{\text{G-LOD}}$ is an H^1-approximation of u_h and $((I_H|_{V_H})^{-1} \circ I_H)(u_H^{\text{G-LOD}})$ an L^2-approximation of u_h, respectively. Furthermore, because $k^{\frac{d}{2}} \theta^k$ converges with exponential order to zero, the error $|||u_h - u_H^{\text{G-LOD}}|||_h$ is typically dominated by the first term of order $\mathrm{O}(H)$. This was observed in various numerical experiments in different works, cf. [19,20,35]. In particular, a specific choice $k \gtrsim (p+1)|\log(H)|$ leads to a $\mathrm{O}(H)$ convergence for the total H^1-error, see also [19,20,35].

3.2 Petrov–Galerkin LOD

In a straightforward manner, we can now state the LOD in Petrov–Galerkin formulation: find $u_H^{\text{PG-LOD}} \in V^{\text{ms}}$ that satisfies

$$a_h \left(u_H^{\text{PG-LOD}}, \Phi_H \right) = (f, \Phi_H) \quad \text{for all } \Phi_H \in V_H. \tag{12}$$

A unique solution of (12) is guaranteed by the inf-sup stability. In practice, inf-sup stability is clearly observable in numerical experiments (see Sect. 4). Analytically we can make the following observations.

Remark 2 (*Quasi-orthogonality and inf-sup stability*) The inf-sup stability of the LOD in Petrov–Galerkin formulation is a natural property to expect, since we have quasi-orthogonality in $a_h(\cdot, \cdot)$ of the spaces V^{ms} and W_h. This can be verified by a simple computation. Let $\Phi^{\text{ms}} = \Phi_H + Q_h(\Phi_H) \in V^{\text{ms}}$, let $w_h \in W_h$ and let $Q_h^{\Omega}(\Phi_H)$ the optimal corrector as in assumption (A8), then

$$\begin{aligned} a_h(\Phi^{\text{ms}}, w_h) &= a_h(\Phi_H + Q_h(\Phi_H), w_h) \\ &= a_h \left(Q_h(\Phi_H) - Q_h^{\Omega}(\Phi_H), w_h \right) \\ &\le \left|\left|\left| Q_h(\Phi_H) - Q_h^{\Omega}(\Phi_H) \right|\right|\right|_h |||w_h|||_h \\ &\lesssim k^{d/2} \theta^k (1/H)^p \left|\left|\left| \Phi_H + Q_h^{\Omega}(\Phi_H) \right|\right|\right|_h |||w_h|||_h, \end{aligned}$$

with generic constants $0 < \theta < 1$ and $p \in \{0, 1\}$ as in (A8). This means that $a_h(\Phi^{\text{ms}}, w_h)$ converges exponentially in k to zero, and it is identical to zero for all sufficiently large k [because then $Q_h(\Phi_H) = Q_h^{\Omega}(\Phi_H)$]. Writing the PG-LOD bilinear form as

$$\begin{aligned}&a_h(\Phi_H + Q_h(\Phi_H), \Psi_H)\\ &\quad = a_h(\Phi_H + Q_h(\Phi_H), \Psi_H + Q_h(\Psi_H)) + a_h(\Phi_H + Q_h(\Phi_H), Q_h(\Psi_H)),\end{aligned}$$

we see that it is only a small perturbation of the symmetric (coercive) G-LOD version, where the difference can be bounded by the quasi-orthogonality.

Even though the quasi-orthogonality *suggests* inf-sup stability, the given assumptions (A1)–(A8) do not seem to be sufficient for rigorously proving it. Here, it seems necessary to leave the abstract setting and to prove the inf-sup stability result for the various LOD realizations separately. For simplification, we therefore make the inf-sup stability to be an additional assumption [see (A9) below]. Later we give an example how to prove this assumption for a certain realization of the method. We also note that the inf-sup stability can be always verified numerically (for a given k) by investigating the system matrix $S^{\text{PG-LOD}}$ given by the entries

$$(S^{\text{PG-LOD}})_{ij} = a_h(\Phi_j + Q_h(\Phi_j), \Phi_i)$$

for $1 \leq i, j \leq N_H$ where N_H denotes the dimension of V_H and where $\{\Phi_i | \, 1 \leq i \leq N_H\}$ denotes a basis of V_H. To check the inf-sup stability we must compute the eigenvalues of $S^{\text{PG-LOD}}$. If their real parts are all strictly positive, we have inf-sup stability and the inf-sup constant is identical to the smallest real part of an eigenvalue. Standard approaches for computing the eigenvalues of a non-symmetric matrix are the Arnoldi method, the Jacobi–Davidson method and the non-symmetric Lanczos algorithm (cf. [39] for a comprehensive overview). Since N_H is moderately small, the cost for applying one of the methods are still feasible.

(A9) We assume that the LOD in Petrov–Galerkin formulation is inf-sup stable in the following sense: there exists a sequence of constants $\alpha(k)$ and a generic limit $\alpha_0 > 0$ (independent of H, h, k or the oscillations of A) such that $\alpha(k)$ converges with *exponential speed* to α_0, i.e. there exist constants $C(H)$ (possibly depending on H, but not on h, k or the oscillations of A) and a generic $\theta \in (0, 1)$ such that $|\alpha(k) - \alpha_0| \leq C(H) k^{d/2} \theta^k$. Furthermore it holds $\alpha(\bar{k}) = \alpha_0$ for all sufficiently large $\bar{k}$ and

$$\frac{a_h(\Phi^{\text{ms}}, \Phi_H)}{|||\Phi_H|||_H} \geq \alpha(k) |||\Phi^{\text{ms}}|||_h,$$

for all $\Phi^{\text{ms}} \in V^{\text{ms}}$ and $\Phi_H := ((I_H|_{V_H})^{-1} \circ I_H)(\Phi^{\text{ms}}) \in V_H$.

The following result states that the approximation quality of the LOD in Petrov–Galerkin formulation is of the same order as for the Galerkin LOD, up to a possible pollution term depending on $C_{H,h}$, but which still converges exponentially to zero.

Theorem 2 (A priori error estimate for PG-LOD) *Assume (A1)–(A9). Given a positive $k \in \mathbb{N}_{>0}$, let for all $T \in \mathcal{T}_H$ the patch $U(T) = U_k(T)$ be defined as in (8) and large* *enough so that the inf-sup constant in (A9) fulfills* $\alpha(k) \geq \bar{\alpha}$ *for some* $\bar{\alpha} > 0$ *and let* $u_H^{\textit{PG-LOD}}$ *be the unique solution of* (12). *Let* $u_h \in V_h$ *be the fine scale reference solution governed by* (2). *Then, the following a priori error estimate holds true*

$$\begin{aligned} &\left\| u_h - ((I_H|_{V_H})^{-1} \circ I_H)(u_H^{\textit{PG-LOD}}) \right\|_{L^2(\Omega)} + |||u_h - u_H^{\textit{PG-LOD}}|||_h \\ &\lesssim (H + (1/H)^p (1 + (1/\bar{\alpha}))(1 + C_{H,h}) k^{d/2} \theta^k) \|f\|_{L^2(\Omega)}, \end{aligned}$$

where $0 < \theta < 1$ *and* $p \in \{0, 1\}$ *are the generic constants from assumption* (*A8*) *and* $C_{H,h}$ *as in* (*A5*).

3.3 Example 1: continuous Galerkin finite element method

The previous subsection showed that the Petrov–Galerkin formulation of the LOD does not suffer from a loss in accuracy with respect to the symmetric formulation. In this subsection, we give the specific example of the LOD for the continuous Galerkin finite element method. In particular, we discuss the advantage of the PG formulation over the symmetric formulation. Let us first introduce the specific setting and the corresponding argument about the validity of (A4)–(A9) on this setting.

In addition to the assumptions that we made on the shape regular partitions $\mathcal{T}_H$ and $\mathcal{T}_h$ in Sect. 2.1, we assume that $\mathcal{T}_H$ and $\mathcal{T}_h$ are either triangular or quadrilateral meshes. Accordingly, for $\mathcal{T} = \mathcal{T}_H, \mathcal{T}_h$ we denote

$$\begin{aligned} P_1(\mathcal{T}) &:= \left\{ v \in C^0(\Omega) | \forall T \in \mathcal{T}, v|_T \text{ is a polynomial of total degree } \leq 1 \right\} \quad \text{and} \\ Q_1(\mathcal{T}) &:= \left\{ v \in C^0(\Omega) | \forall T \in \mathcal{T}, v|_T \text{ is a polynomial of partial degree } \leq 1 \right\} \end{aligned}$$

and define $V_h := P_1(\mathcal{T}_h) \cap H_0^1(\Omega)$ if $\mathcal{T}_h$ is simplicial and $V_h := Q_1(\mathcal{T}_h) \cap H_0^1(\Omega)$ if it is a quadrilation. The coarse space $V_H \subset V_h$ is defined in the same fashion and since $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$, assumption (A6) is obviously fulfilled. For simplicity, we also assume that the coarse mesh $\mathcal{T}_H$ is quasi-uniform (which is the typical choice in applications).

The bilinear form $a_h(\cdot, \cdot)$ is defined by the standard energy scalar product on $H_0^1(\Omega)$ that belongs to the elliptic problem to solve, i.e.

$$a_h(v, w) := \int_\Omega A \nabla v \cdot \nabla w \quad \text{for } v, w \in H_0^1(\Omega).$$

Accordingly, we set $|||v|||_h := |||v|||_H := \|A^{1/2} \nabla v\|_{L^2(\Omega)}$ for $v \in H^1(\Omega)$. Hence, assumptions (A5) and (A6) are fulfilled and the solution $u_h \in V_h$ of (2) is nothing but the standard continuous Galerkin finite element solution on the fine grid $\mathcal{T}_h$.

Next, we specify $I_H \colon V_h \to W_h$ in (A7). For this purpose, let $\Phi_z \in V_H$ be the nodal basis function associated with the coarse grid node $z \in \mathcal{N}_H$, i.e., $\Phi_z(y) = \delta_{yz}$. Let I_H be the weighted Clément-type quasi-interpolation operator as defined in [9,10]:

$$I_H\colon H_0^1(\Omega) \rightarrow V_H, \quad v \mapsto I_H(v) := \sum_{z \in \mathcal{N}_H^0} v_z \Phi_z \quad \text{with } v_z := \frac{(v, \Phi_z)_{L^2(\Omega)}}{(1, \Phi_z)_{L^2(\Omega)}}. \tag{13}$$

First we note that it was shown in [35] that $(I_H)_{|V_H}\colon V_H \rightarrow V_H$ is an isomorphism [but not a projection, i.e. $(I_H|_{V_H})^{-1} \neq I_H|_{V_H}$]. Hence, $(I_H)_{|V_H}^{-1}$ exists. This is one of the properties in (A7). The L^2- and H^1-stability of I_H, as well as corresponding approximation properties, were proved in [9]. It only remains to check the H^1-stability of $(I_H)_{|V_H}^{-1}$. Unfortunately, this property is not trivial to fulfill. First, we note that it was shown in [34] that the mapping $(I_H)_{|V_H}^{-1} \circ I_H$ is nothing but the L^2-projection $P_{L^2}\colon H_0^1(\Omega) \rightarrow V_H$ (see also Remark 5 below). Consequently, the question of H^1-stability of $(I_H)_{|V_H}^{-1}$ is equivalent to the question of H^1-stability of the L^2-projection. This result is well-established for quasi uniform grids (cf. [6]) as assumed at the beginning of this section. However it is still open for arbitrary refinements. The most recent results on this issue can be found in [7,15,29], where the desired H^1-stability was shown for certain types of adaptively refined meshes. To avoid complicated mesh assumptions in this paper, we simply assume $\mathcal{T}_H$ to be quasi-uniform. This is not very restrictive since adaptive refinements should typically take place on the fine mesh $\mathcal{T}_h$. Alternatively, in light of [7,15,29], we could also directly assume that the L^2-projection on V_H is H^1-stable to allow more general coarse meshes.

It remains to specify $a_h^T(\cdot, \cdot)$, which we define by

$$a_h^T(v, w) := \int_T A\nabla v \cdot \nabla w \quad \text{for } v, w \in H_0^1(\Omega).$$

Let us for simplicity denote $|||\cdot|||_{h,T} := \|A^{1/2}\nabla \cdot\|_{L^2(T)}$. The decay assumption (A8) was essentially proved in [19, Lemma 3.6], which established the existence of a generic constant $0 < \theta < 1$ with the properties as in (A8) such that

$$|||(Q_h - Q_h^{\Omega})(\Phi_H)|||_h^2 \lesssim k^d \theta^{2k} \sum_{T \in \mathcal{T}_H} \left|\left|\left| Q_h^{\Omega,T}(\Phi_H) \right|\right|\right|_h^2, \tag{14}$$

for all $\Phi_H \in V_H$. On the other hand we have by $|||\cdot|||_{h,T} = \|A^{1/2}\nabla \cdot\|_{L^2(T)}$ and Eq. (6) that

$$\begin{aligned} \left|\left|\left| Q_h^{\Omega,T}(\Phi_H) \right|\right|\right|_h^2 &\lesssim a_h\left(Q_h^{\Omega,T}(\Phi_H), Q_h^{\Omega,T}(\Phi_H)\right) \\ &= -a_h^T\left(\Phi_H, Q_h^{\Omega,T}(\Phi_H)\right) \\ &\lesssim |||\Phi_H|||_{h,T} \left|\left|\left| Q_h^{\Omega,T}(\Phi_H) \right|\right|\right|_h. \end{aligned} \tag{15}$$

Hence, by plugging this result into (14):

$$\begin{aligned}\left|\left|\left|(Q_h - Q_h^{\Omega})(\Phi_H)\right|\right|\right|_h^2 &\lesssim k^d \theta^{2k} \sum_{T \in \mathcal{T}_H} |||\Phi_H|||_{h,T}^2 \\ &\lesssim k^d \theta^{2k} |||\Phi_H|||_h^2 = k^d \theta^{2k} |||((I_H|_{V_H})^{-1} \circ I_H)(\Phi_H + Q_h^{\Omega}(\Phi_H))|||_h^2 \\ &\overset{(A7)}{\lesssim} k^d \theta^{2k} |||\Phi_H + Q_h^{\Omega}(\Phi_H)|||_h^2,\end{aligned}$$

which proves that assumption (A8) holds even with $p = 0$. The remaining assumption (A9) is less obvious and requires a proof. We give this proof for the continuous Galerkin PG-LOD in Sect. 5. We summarize the result in the following lemma.

Lemma 1 (Inf-sup stability of continuous Galerkin PG-LOD) *For all* $T \in \mathcal{T}_H$ *let* $U(T) = U_k(T)$ *for* $k \in \mathbb{N}$. *Then there exist generic constants* C_1, C_2 *(independent of* H, h, k *or the oscillations of* A*) and* $0 < \theta < 1$ *as in assumption (A8), so that it holds*

$$\inf_{\Phi_H \in V_H} \sup_{\Phi^{\mathrm{ms}} \in V^{\mathrm{ms}}} \frac{a(\Phi^{\mathrm{ms}}, \Phi_H)}{|||\Phi^{\mathrm{ms}}|||_h |||\Phi_H|||_h} \geq \alpha(k),$$

for $\alpha(k) := C_1 \alpha - C_2 k \theta^k \omega(\Phi^{\mathrm{ms}})$ *and*

$$0 \leq \omega(\Phi^{\mathrm{ms}}) := \inf_{w_h \in W_h^T} \frac{\left\| \nabla \Phi^{\mathrm{ms}} - \nabla((I_H|_{V_H})^{-1} \circ I_H)(\Phi^{\mathrm{ms}}) - \nabla w_h \right\|}{\left\| \nabla \Phi^{\mathrm{ms}} - \nabla((I_H|_{V_H})^{-1} \circ I_H)(\Phi^{\mathrm{ms}}) \right\|} \leq 1,$$

where $W_h^T := \{w_h \in W_h |\, w_h|_T \in W_h(T)\}$, *i.e. the space of all functions from* W_h *that are zero on the boundary of the coarse grid elements. Observe that* $\alpha(k)$ *converges with exponential speed to* αC_1. *Furthermore we have* $\alpha(0) = C_1 \alpha$ *[because* $\omega(\Phi^{\mathrm{ms}}) = 0$*] and also* $\alpha(\ell) = C_1 \alpha$ *for all sufficiently large* ℓ.

Remark 3 Let $U(T) = U_k(T)$ for $k \in \mathbb{N}$ with $k \gtrsim |\log(H)|$, then the CG-LOD in Petrov–Galerkin formulation is inf-sup stable for sufficiently small H. In particular, there exists a unique solution of problem (12).

Remark 4 Lemma 1 does not allow to conclude to inf-sup stability for the regime $0 < k \ll |\log(H)|$. However, even though this regime is not of practical relevance, it is interesting to note that we could not observe a violation of the inf-sup stability for any value of k and in any numerical experiment that we set up so far.

Since assumptions (A1)–(A9) are fulfilled for this setting, Theorems 1 and 2 hold true for the arising method. Furthermore, we have $p = 0$ and $C_{H,h} = 1$ in the estimates, meaning that the $(1/H)$-pollution in front of the decay term vanishes. We can summarize the result in the following conclusion.

Conclusion 3 *Assume the (continuous Galerkin) setting of this subsection and let* $u_H^{\text{PG-LOD}}$ *denote a Petrov–Galerkin solution of* (12). *If* $k \gtrsim mH|\log(H)|$ *for* $m \in \mathbb{N}$, *then it holds*

$$\left\| u_h - u_H^{\text{PG-LOD}} \right\|_{H^1(\Omega)} \lesssim (H + H^m) \|f\|_{L^2(\Omega)}.$$

In particular, the bound is independent of $C_{H,h}$.

3.4 Discussion of advantages

The central disadvantage of the Galerkin LOD is that it requires a communication between solutions of different patches. Consider for instance the assembly of the system matrix that belongs to problem (10). Here it is necessary to compute entries of the type

$$\int_{\Omega} A\nabla(\Phi_i + Q_h(\Phi_i)) \cdot \nabla(\Phi_j + Q_h(\Phi_j)),$$

which particularly involves the computation of the term

$$\sum_{\substack{T\in\mathcal{T}_H \\ T\subset\omega_i}} \sum_{\substack{K\in\mathcal{T}_H \\ K\subset\omega_j}} \int_{U(T)\cap U(K)} A\nabla Q_h^T(\Phi_i) \cdot \nabla Q_h^K(\Phi_j), \tag{16}$$

where $\Phi_i, \Phi_j \in V_H$ denote two coarse nodal basis functions and ω_i and ω_j its corresponding supports. The efficient computation of (16) requires information about the intersection area of any two patches $U(T)$ and $U(K)$. Even if T and K are not adjacent or close to each other, the intersection of the corresponding patches can be complicated and non-empty. The drawback becomes obvious: first, these intersection areas must be determined, stored and handled in an efficient way and second, the number of relevant entries of the stiffness matrix (i.e. the non-zeros) increases considerably. Note that this also leads to a restriction in the parallelization capabilities, in the sense that the assembly of the stiffness matrix can only be 'started' if the correctors $Q_h(\Phi_i)$ are already computed. Another disadvantage is that the assembly of the right hand side vector associated with (f, Φ^{ms}) in (10) is much more expensive since it involves the computation of entries $(f, \Phi_i + Q_h(\Phi_i))_{L^2(\Omega)}$. First, the integration area is $\cup\{U(T)|\, T \in \mathcal{T}_H,\ T \subset \omega_i\}$ instead of typically ω_i. This increases the computational costs. At the same time, it is also hard to assemble these entries by performing (typically more efficient) element-wise computations (for which each coarse element has to be visited only once). Second, $(f, \Phi_i + Q_h(\Phi_i))_{L^2(\Omega)}$ involves a quadrature rule of high order, since $Q_h(\Phi_i)$ is rapidly oscillating. These oscillations must be resolved by the quadrature rule, even if f is a purely macroscopic function that can be handled exactly by a low order quadrature. Hence, the costs for computing $(f, \Phi_i + Q_h(\Phi_i))_{L^2(\Omega)}$ depend indirectly on the oscillations of A. Finally, if the LOD shall be applied to a sequence of problems of type (1), which only differ in the source term f (or a boundary condition), the system matrix can be fully reused, but the complications that come with the right hand side have to be addressed each time again.

The Petrov–Galerkin formulation of the LOD clearly solves these problems without suffering from a loss in accuracy. In particular:

- The PG-LOD does not require any communication between two different patches and the resulting stiffness matrix is sparser than the one for the symmetric LOD. In particular, the entries of the system matrix S can be computed with the following algorithm:

```
Let $S$ denote the empty system matrix with entries $S_{ij}$.

Algorithm: assembleSystemMatrix( $\mathcal{T}_H, \mathcal{T}_h, k$ )

In parallel **foreach** $T \in \mathcal{T}_H$ **do**
    **foreach** $z_i \in \mathcal{N}_H^0$ with $z_i \in \overline{T}$ **do**
        compute $Q_h^T(\Phi_{z_i}) \in W_h(U_k(T))$ with
            $a(Q_h^T(\Phi_{z_i}), w_h) = -\int_T A\nabla\Phi_{z_i} \cdot \nabla w_h \quad \text{for all } w_h \in W_h(U_k(T)).$
        **foreach** $z_j \in \mathcal{N}_H^0$ with $z_j \in \overline{U(T)}$ **do**
            update the system matrix:
                $S_{ji} \mathrel{+}= \int_{\omega_j} A\left(\Phi_{z_i} + \nabla Q_h^T(\Phi_{z_i})\right) \cdot \nabla\Phi_{z_j}.$
        **end**
    **end**
**end**
```

  Observe that it is possible to add the local terms $a(\Phi_{z_i} + Q_h^T(\Phi_{z_i}), \Phi_{z_j})$ directly to the system matrix S, i.e. the assembling of the matrix is parallelized in a straightforward way and does not rely on the availability of other results.
- Replacing the source term f in (1), only involves the re-computation of the terms $(f, \Phi_i)_{L^2(\omega_i)}$ for coarse nodal basis functions Φ_i, i.e. the same costs as for the standard FE method on the coarse scale. Furthermore, the choice of the quadrature rule relies purely on f, but not on the oscillations of A.

Besides the previously mentioned advantages, there is still a memory consuming issue left: the storage of the local correctors $Q_h^T(\Phi_{z_i})$. These local correctors need to be saved in order to express the final approximation $u_H^{\text{PG-LOD}}$ which is spanned by the multiscale basis functions $\Phi_i + Q_h(\Phi_i)$. As long as we are interested in a good H^1-approximation of the solution, this problem seems to be unavoidable. However, in many applications we can even overcome this difficulty by exploiting another very big advantage of the PG-LOD: Theorem 2 predicts that alone the 'coarse part' of $u_H^{\text{PG-LOD}}$, denoted by $u_H := ((I_H|_{V_H})^{-1} \circ I_H)(u_H^{\text{PG-LOD}})) \in V_H$, already exhibits very good L^2-approximation properties, i.e. if $k \gtrsim |\log(H)|$ we have essentially

$$\|u_h - u_H\|_{L^2(\Omega)} \leq \mathrm{O}(H).$$

In contrast to $u_H^{\text{PG-LOD}}$, the representation of u_H does only require the classical coarse finite element basis functions. Hence, we can use the algorithm presented earlier, with the difference that we can immediately delete $Q_h^T(\Phi_i)$ after updating the stiffness matrix. Observe that even if computations have to be repeated for different source terms f, this stiffness matrix can be reused again and again. Also, if a user is interested in the fine scale behavior in a local region [but the $Q_h^T(\Phi_i)$ were already dropped], it is still possible to quickly re-compute the desired local corrector for the region.

As an application, consider for instance the case that the problem

$$\int_\Omega A\nabla u \cdot \nabla v = \int_\Omega f v$$

describes the diffusion of a pollutant in groundwater. Here, u describes the concentration of the pollutant, A the (rapidly varying) hydraulic conductivity and f a source term describing the injection of the pollutant. In such a scenario, there is typically not much interest in finding a good approximation of the (locally fluctuating) gradient ∇u, but rather in the macroscopic behavior of pollutant u, i.e. in purely finding a good L^2-approximation that allows to conclude where the pollutant spreads. A similar scenario is the investigation of the properties of a composite material, where A describes the heterogenous material and f some external force. Again, the interest is in finding an accurate L^2-approximation. Besides, the corresponding simulations are typically performed for a variety of different source terms f, investigating different scenarios. In this case, the PG-LOD yields reliable approximations with very low costs, independent of the structure of A.

Remark 5 (*Relation to the L^2-projection*) Assume the setting of this subsection. In [34] it was shown that $(v_H, w_h)_{L^2(\Omega)} = 0$ for all $v_H \in V_H$ and $w_h \in W_h$, i.e. V_H and W_h are L^2-orthogonal. This implies that

$$(I_H|_{V_H})^{-1} \circ I_H = P_{L^2},$$

with P_{L^2} denoting the L^2-projection on V_H. To verify this, let $v_h \in V_h$ be arbitrary. Then due to $V_h = V_H \oplus W_h$ we can write $v_h = v_H + w_h$ (with $v_H \in V_H$ and $w_h \in W_h$) and observe for all $\Phi_H \in V_H$

$$\begin{aligned}\int_\Omega P_{L^2}(v_h) \quad \Phi_H &= \int_\Omega v_h \quad \Phi_H \overset{V_H \perp_{L^2} W_h}{=} \int_\Omega v_H \quad \Phi_H \\ &= \int_\Omega ((I_H|_{V_H})^{-1} \circ I_H)(v_H) \quad \Phi_H \overset{I_H(w_h)=0}{=} \int_\Omega ((I_H|_{V_H})^{-1} \circ I_H)(v_h) \quad \Phi_H.\end{aligned}$$

Hence, $u_H^{\text{PG-LOD}} = u_H + Q_h(u_H)$ with $u_H = P_{L^2}(u_H^{\text{PG-LOD}})$.

Conclusion 4 (Application to homogenization problems) *Assume the setting of this subsection and let P_{L^2} denote the L^2-projection on V_H as in Remark 5. We consider now a typical homogenization setting with $(\epsilon)_{>0} \subset \mathbb{R}_{>0}$ being a sequence of positive parameters that converges to zero. Let $Y := [0,1]^d$ denote the unique cube in $\mathbb{R}^d$ and let $A^\epsilon(x) = A_p(x, \frac{x}{\epsilon})$ for a function $A_p \in W^{1,\infty}(\Omega \times Y)$ that is Y-periodic in the second argument (hence A^ϵ is rapidly oscillating with frequency ϵ). The corresponding exact solution of problem* (1) *shall be denoted by $u_\epsilon \in H_0^1(\Omega)$. It is well known* (*cf.* [3])

that u_ϵ converges weakly in H^1 (but not strongly) to some unique function $u_0 \in H_0^1(\Omega)$. Furthermore, if $\|f\|_{L^2(\Omega)} \lesssim 1$ it holds $\|u_\epsilon - u_0\|_{L^2(\Omega)} \lesssim \epsilon$. With Theorem 2 together with Remark 5 and standard error estimates for FE problems, we hence obtain:

$$\|u_0 - u_H\|_{L^2(\Omega)} \lesssim \epsilon + \left(\frac{h}{\epsilon}\right)^2 + H,$$

for $u_H = P_{L^2}(u_H^{\text{PG-LOD}})$. Homogenization problems are typical problems, where one is often purely interested in the L^2-approximation of the exact solution u_ϵ, meaning one is interested in the homogenized solution u_0.

As discussed in this section, the PG-LOD can have significant advantages over the (symmetric) G-LOD with respect to computational costs, efficiency and memory demand. In Sect. 4.1 we additionally present a numerical experiment to demonstrate that the approximations produced by the PG-LOD are in fact very close to the ones produced by (symmetric) G-LOD, i.e. not only of the same order as predicted by the theorems, but also of the same quality.

Remark 6 (*Nonlinear problems*) The above results suggest that the advantages can become even more pronounced for certain types of nonlinear problems. For instance, consider a well-posed problem of the type

$$-\nabla \cdot A\nabla u + c(u) = f,$$

for a nonlinear function c. Here, it is intuitively reasonable to construct $Q_h(\Phi_H)$ as before using only the linear elliptic part of the problem. This is a preprocessing step that is done once and can be immediately deleted stiffness matrix is calculated and saved. Then we solve for $u_H \in V_H$ that satisfies

$$(A\nabla(u_H + Q_h(u_H)), \nabla\Phi_H)_{L^2(\Omega)} + (c(u_H), \Phi_H)_{L^2(\Omega)} = (f, \Phi_H)_{L^2(\Omega)}$$

for all $\Phi_H \in V_H$. Clearly, typical iterative solvers can be utilized to solve this variational problem. This iteration is inexpensive because it is done in V_H and the pre-constructed stiffness matrix can be fully reused within every iteration and since the other contributions are independent of Q_h. Performing iterations on the coarse space for solving nonlinear problems within the framework of multiscale finite element has been investigated (see for example [12,16]).

3.5 Example 2: discontinuous Galerkin finite element method

In this subsection, we apply the results of Sect. 3.2 to a LOD Method that is based on a discontinuous Galerkin approach. The DG-LOD was originally proposed in [14] and fits into the framework proposed in Sect. 2.2. First, we show that the setting fulfills assumptions (A4)–(A8) and after we discuss the advantage of the PG DG-LOD over the symmetric DG-LOD. For simplification, we assume that A is piecewise constant with respect to the fine mesh $\mathcal{T}_h$ so that all of the subsequent traces are well-defined.

Again, we make the same assumptions on the partitions $\mathcal{T}_H$ and $\mathcal{T}_h$ as in Sect. 2.1 and additionally assume that $\mathcal{T}_H$ and $\mathcal{T}_h$ are either triangular or quadrilateral meshes. The corresponding total sets of edges (or faces for $d = 3$) are denoted by $\mathcal{E}_h$ (for $\mathcal{T}_h$), where $\mathcal{E}_h(\Omega)$ and $\mathcal{E}_h(\partial\Omega)$ denotes the set of interior and boundary edges, respectively.

Furthermore, for $\mathcal{T} = \mathcal{T}_H, \mathcal{T}_h$ we denote the spaces of discontinuous functions with total, respectively partial, polynomial degree equal to or less than 1 by

$$\mathcal{P}_1(\mathcal{T}) := \left\{v \in L^2\Omega)|\forall T \in \mathcal{T}, v|_T \text{ is a polynomial of total degree} \le 1\right\} \quad \text{and}$$
$$\mathcal{Q}_1(\mathcal{T}) := \left\{v \in L^2(\Omega)|\forall T \in \mathcal{T}, v|_T \text{ is a polynomial of partial degree} \le 1\right\}$$

and define $V_h := P_1(\mathcal{T}_h)$ if $\mathcal{T}_h$ is a triangulation and $V_h := Q_1(\mathcal{T}_h)$ if it is a quadrilation. The coarse space $V_H \subset V_h$ is defined in the same fashion with $\mathcal{T}_H$ instead of $\mathcal{T}_h$. Note that these spaces are no subspaces of $H^1(\Omega)$ as in the previous example. For this purpose, we define ∇_h to be the $\mathcal{T}_h$-piecewise gradient [i.e. $(\nabla_h v_h)|t := \nabla(v_h|t)$ for $v_h \in V_h$ and $t \in \mathcal{T}_h$].

For every edge/face $e \in \mathcal{E}_h(\Omega)$ there are two adjacent elements $t^-, t^+ \in \mathcal{T}_h$ with $e = \partial t^- \cap \partial t^+$. We define the jump and average operators across $e \in \mathcal{E}_h(\Omega)$ by

$$[v] := (v|t^- - v|t^+) \quad \text{and} \quad \{A\nabla v \cdot n\} := \frac{1}{2}((A\nabla v)|t^- + (A\nabla v)|t^+) \cdot n,$$

where n be the unit normal on e that points from t^- to t^+, and on $e \in \mathcal{E}_h(\partial\Omega)$ by

$$[v] := w|t \quad \text{and} \quad \{A\nabla v \cdot n\} := (A\nabla v)|t \cdot n$$

where n is the outwards unit normal of $t \in \mathcal{T}_h$ (and Ω). Observe that flipping the roles of t^- and t^+ leads to the same terms in the bilinear form defined below.

With that, we can define the typical bilinear form that characterizes the discontinuous Galerkin method:

$$\begin{aligned} a_h(v_h, w_h) := (A\nabla_h v_h, \nabla_h w_h)_{L^2(\Omega)} + \sum_{e\in\mathcal{E}_h} \frac{\sigma}{h_e}([v_h], [w_h])_{L^2(e)} \\ - \sum_{e\in\mathcal{E}_h} \left((\{A\nabla v_h \cdot n\}, [w_h])_{L^2(e)} + (\{A\nabla w_h \cdot n\}, [v_h])_{L^2(e)}\right). \end{aligned}$$

Here, σ is a penalty parameter that is chosen sufficiently large and $h_e = \text{diam}(e)$. The coarse bilinear form $a_H(\cdot,\cdot)$ is defined analogously with coarse scale quantities. It is well known, that $a_h(\cdot,\cdot)$ [respectively $a_H(\cdot,\cdot)$] is a scalar product on V_h (respectively V_H). Consequently (A4) is fulfilled. As a norm on V_h that fulfills assumption (A5), we can pick

$$|||v|||_h := \left\| A^{1/2} \nabla_h v \right\|_{L^2(\Omega)} + \left(\sum_{e \in \mathcal{E}_h} \frac{\sigma}{h_e} \|[v]\|^2_{L^2(e)} \right)^{1/2}.$$

Analogously, we define $|||v|||_H$ to be a norm on V_H. In this case we obtain the constant $C_{H,h} = \sqrt{H/h}$. Assumption (A6) is obviously fulfilled.

As the operator in assumption (A7) we pick the L^2-projection on V_H, i.e. for $v_h \in V_h$ we have

$$(I_h(v_h), \Phi_H)_{L^2(\Omega)} = (v_h, \Phi_H)_{L^2(\Omega)} \quad \text{for all } \Phi_H \in V_H.$$

In [14, Lemma 5] it was proved that the operator fulfills the desired approximation and stability properties. Since I_H is a projection, we have $I_H = (I_H|_{V_H})^{-1}$ and hence obviously also $|||\cdot|||_H$-stability of the inverse on V_H.

The localized bilinear form $a_h^T(\cdot,\cdot)$ in (5) is defined by $a_h^T(v_h, w_h) := a_h(\chi_T v_h, w_h)$ where $\chi_T = 1$ in T and 0 otherwise, is the element indicator function. Obviously we have for all $v_h, w_h \in V_h$ that

$$a_h(v_h, w_h) = \sum_{T \in \mathcal{T}_H} a_h^T(v_h, w_h).$$

In [14] the DG-LOD is presented in a slightly different way, in the sense that there exists no general corrector operator Q_h. Instead, 'basis function correctors' are introduced. However, it is easily checkable that each of these 'basis function correctors' is nothing but the corrector operator, defined via (6), applied to an original coarse basis function. Therefore, the correctors given by (6) are just an extension of the definition to arbitrary coarse functions. Hence, both methods coincide and are just presented in a different way.

Next, we discuss (A8). This property was shown in [14, Lemmas 11 and 12], however not explicitly for the setting that we established in Definition 3. It was only shown for $\Phi_H = \lambda_{T,j}$, where $\lambda_{T,j} \in V_H$ denotes a basis function on T associated with the j'th node. However, the proofs in [14] directly generalize to the local correctors $Q_h^T(\Phi_H)$ given by Eq. (6). More precisely, following the proofs in [14] it becomes evident that the availability of the required decay property (A8) purely relies on the fact, that the right hand side in the local problems is only locally supported (with a support that remains fixed, even if the patch size decreases). Therefore (A8) can be proved analogously.

Finally, assumption (A9) is not easy to verify. It is obviously fulfilled for the case $U(T) = \Omega$, but the generalized result is harder to verify. The following result holds under some restrictions on the meshes $\mathcal{T}_H$ and $\mathcal{T}_h$.

Lemma 2 (Inf-sup stability of discontinuous Galerkin PG-LOD) *Assume that $\mathcal{T}_H$ is quasi-uniform and that there exists an exponent $m \in \mathbb{R}$ with $m > 1$ such that for all $T \in \mathcal{T}_H$*

$$\operatorname{diam}(T)^m \lesssim \min \left\{ h_e | \, e \in \mathcal{E}_h \text{ and } e \subset \overline{T} \right\}$$

(i.e. if $\mathcal{T}_h$ is also quasi-uniform we assume $H^m \lesssim h$). If $k \in \mathbb{N}$ is such that $k \gtrsim \frac{(m+3)}{2}|\log(H)|$ then, for sufficiently small H, there exist generic positive constants C_1, C_2 such that

$$\inf_{\Phi_H \in V_H} \sup_{\Phi^{\mathrm{ms}} \in V^{\mathrm{ms}}} \frac{a_h(\Phi^{\mathrm{ms}}, \Phi_H)}{|||\Phi^{\mathrm{ms}}|||_h |||\Phi_H|||_H} \geq C_1(\alpha - C_2 H).$$

Hence, we have inf-sup stability for sufficiently small H.

The proof is given in Sect. 5. We note that the inf-sup stability can be observed numerically already under weaker assumptions (see Sect. 4) and that it is in general ‘a reasonable thing to expect’ as discussed in Remark 2.

In conclusion, the discontinuous Galerkin LOD in Petrov–Galerkin formulation fulfills the assumptions of our framework [up to a discussion on (A9)]. The advantages that we discussed in the previous subsection for the Petrov–Galerkin continuous finite element method in terms of memory and efficiency remains true. However, for the PG DG-LOD there is a very important additional advantage. It is known that the classical DG method has the feature of local mass conservation with respect to the elements of the underlying mesh. This can be easily checked by testing with the indicator function of an element T in the variational formulation of the method. The local mass conservation is a highly desired property for various flow and transport problems. However, the DG-LOD does not preserve this property, since the indicator function of an element (whether coarse or fine) is not in the space V^{ms}. This problem is solved in the PG DG-LOD, where we can test with any element from V_H and in particular with the indicator function of a coarse element. Hence, in contrast to the symmetric DG-LOD, the PG DG-LOD is locally mass conservative with respect to coarse elements $T \in \mathcal{T}_H$. This allows for example the coupling of the PG DG-LOD for an elliptic problem with the solver for a hyperbolic conservation law, which was not possible before without relinquishing the mass conservation. We discuss this further in the next subsection.

3.6 Perspectives towards two-phase flow

In this subsection, we investigate an application of the Petrov–Galerkin DG-LOD in the simulation of two-phase flow as governed by the Buckley–Leverett equation. Specifically, the LOD framework is utilized to solve the pressure equation, which is an elliptic boundary value problem, and is coupled with a solver for a hyperbolic conservation law. The Buckley–Leverett equation can be used to model two-phase flow in a porous medium. Generally, the flow of two immiscible and incompressible fluids is driven by the law of mass balance for the two fluids:

$$\Theta \partial_t S_\alpha + \nabla \cdot \boldsymbol{v}_\alpha = q_\alpha \quad \text{in } \Omega \times (0, T_{end}] \quad \text{for } \alpha = w, n. \tag{17}$$

Here, Ω is a computational domain, $(0, T_{end}]$ a time interval, the unknowns $S_w, S_n : \Omega \rightarrow [0, 1]$ describe the saturations of a wetting and a non-wetting fluid and $\boldsymbol{v}_w$ and $\boldsymbol{v}_n$ are the corresponding fluxes. Furthermore, Θ describes the porosity and q_w and q_n

are two source terms. Darcy's law relates the fluxes with the two unknown pressures p_n and p_w by

$$\boldsymbol{v}_\alpha = -K\frac{k_\alpha(S_\alpha)}{\mu_\alpha}(\nabla p_\alpha - \rho_\alpha \boldsymbol{g}) \quad \text{for } \alpha = w, n.$$

Here, K denotes the hydraulic conductivity, k_w and k_n the relative permeabilities depending on the saturations, μ_w and μ_n the viscosities, ρ_w and ρ_n the densities and $\boldsymbol{g}$ the gravity vector. The saturations are coupled via $S_n+S_w = 1$ and a relation between the two pressures is typically given by the capillary pressure relation $P_c(S_w) = p_n - p_w$ for a monotonically decreasing capillary pressure curve P_c. In this case, we obtain the full two-phase flow system, which consists of two strongly coupled, possibly degenerate parabolic equations. However, if we neglect the gravity and the capillary pressure [i.e. assume that $P_c(S_w) = 0$], the system reduces to the so called Buckley–Leverett system with an elliptic pressure equation and an hyperbolic equation for the saturation:

$$-\nabla \cdot (K\lambda(S)\nabla p) = q \quad \text{and} \quad \Theta\partial_t S + \nabla \cdot (f(S)\boldsymbol{v}) = q_w, \tag{18}$$

where we have $S = S_w$, $p = p_w = p_n$, the total mobility $\lambda(S) := \frac{k_w(S)}{\mu_w} + \frac{k_n(1-S)}{\mu_n} > 0$, the flux $\boldsymbol{v} := -K\lambda(S)\nabla p$ and the flux function $f(S) := \frac{k_w(S)}{\mu_w \lambda(S)}$. The total source is given by $q := \frac{q_w+q_n}{2}$. Observe that (18) is obtained from (17) by summing up the equations for the saturations, using $\partial_t(s_n + s_w) = \partial_t 1 = 0$.

An application for which neglecting the capillary pressure is typically justified are oil recovery processes. Here, a replacement fluid, such as water or liquid carbon dioxide, is injected with very high rates into a reservoir to move oil towards a production well. However, often oil is trapped at interfaces of a low and a high conductivity region. This oil would become inaccessible which is why detailed simulations are required before the replacement fluid can be actually injected.

Depending on the choice for the mobilities, the hyperbolic Buckley–Leverett problem can have one or more weak solutions (cf. [32]). One approach for solving the problem numerically is to use an operator splitting technique as proposed in [4], which is more well-known as the *implicit pressure explicit saturation* (IMPES). Here, the hyperbolic Buckley–Leverett problem is treated with an explicit time stepping method where the flux velocity $\boldsymbol{v}$ is kept constant for a certain time interval and then updated by solving the elliptic problem with the saturation from the previous time step (see Fig. 1 for an illustration). Alternatively, depending on the type of the flux function f, the hyperbolic problem can be also solved implicitly with a suitable numerical scheme for conservation laws (cf. [30]) where the flux $\boldsymbol{v}$ arising from the Darcy equation is, as in the previous case, only updated every fixed number of time steps.

Observe that the difficulties produced by the multiscale character of the problem are primarily related to the elliptic part of the problem. Once the Darcy problem is solved to update the flux velocity, the grid for solving the hyperbolic problem can be significantly coarsened. The reason is that $\boldsymbol{v} = -K\lambda(S)\nabla p$ is possibly still rapidly

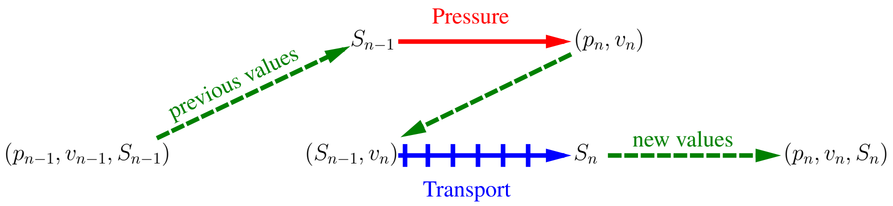

Fig. 1 A schematic of operator splitting (IMPES) for system (18)

oscillating, but the relative amplitude of the oscillations is expected to remain small. In other words, just like for standard elliptic homogenization problems, $\boldsymbol{v}$ behaves like an upscaled quantity $-K_0\lambda(S_0)\nabla p_0$ with effective/homogenized functions K_0, S_0 and p_0.

Remark 7 Any realization of the LOD involves to solve a number of local problems that help us to construct the low dimensional space V^{ms}. One might consider to update this space every time that the Darcy problem has to be solved with a new saturation. However, since $\lambda(S)$ is essentially macroscopic, it is generally sufficient to construct the space only once for $\lambda = 1$ and reuse the result for every time step. This makes solving the elliptic multiscale problem much cheaper after the multiscale space is assembled. A justification for this reusing of the basis can be e.g. found in [20] where it was shown that oscillations coming from advective terms can be often neglected in the construction of a multiscale basis. Under certain assumptions, the relative permeability $\lambda(S)$ can in fact be interpreted as a pure enforcement by an additional advection term.

4 Numerical experiments

In this section we present two different model problems. The first one involves a LOD methods for the continuous Galerkin method. Here, we compare the results obtained with the symmetric version of the method with the results obtained for the Petrov–Galerkin version. In the second model problem, we use a PG DG-LOD for solving the Buckley–Leverett system.

4.1 Continuous Galerkin PG-LOD for elliptic multiscale problems

In this section, we use the setting established in Sect. 3.3. All experiments were performed with the G-LOD and PG-LOD for the continuous finite element method.

In order to be more flexible in the choice of the localization patches $U(T)$, we make subsequently use of “half” or “quarter coarse layers”, i.e. $k \in \mathbb{Q}_{\geq 0}$. This can be easily accomplished by extending Definition 4 straightforwardly to fine grid layers, i.e. for $k \in \mathbb{Q}_{\geq 0}$ and $T \in \mathcal{T}_H$ we define the number of fine layers by $\ell := \lfloor \frac{kH}{h} \rfloor \in \mathbb{N}$ and the corresponding (broken layer) patch by $U_k(T) := U_{\mathrm{f},\ell}(T)$, where iteratively

$U_{\mathrm{f},\ell}(T) := \overline{\cup\{t \in \mathcal{T}_h \mid t \cap U_{\mathrm{f},\ell-1}(T) \neq \emptyset\}}$ and $U_{\mathrm{f},0}(T) := \overline{T}$. This allows us a more careful investigation of the decay behavior.

Let u_h be the solution of (2). In the following we denote by $\|\cdot\|^{\mathrm{rel}}_{L^2(\Omega)}$ and $\|\cdot\|^{\mathrm{rel}}_{H^1(\Omega)}$ the corresponding relative error norms defined by

$$\|u_h - v_h\|^{\mathrm{rel}}_{L^2(\Omega)} := \frac{\|u_h - v_h\|_{L^2(\Omega)}}{\|u_h\|_{L^2(\Omega)}} \quad \text{and}$$
$$\|u_h - v_h\|^{\mathrm{rel}}_{H^1(\Omega)} := \frac{\|u_h - v_h\|_{H^1(\Omega)}}{\|u_h\|_{H^1(\Omega)}}$$

for any $v_h \in V_h$. The coarse part ('the V_H-part') of an LOD approximation $u^{\text{ G-LOD}}$ (respectively $u^{\text{PG-LOD}}$) is subsequently denoted by $P_{L^2}(u^{\text{G-LOD}})$ [respectively $P_{L^2}(u^{\text{ PG-LOD}})$], where P_{L^2} denotes the L^2-projection on V_H (see also Remark 5).

We consider the following model problem. Let $\Omega :=]0,1[^2$ and $\varepsilon := 0.05$. Find $u_\varepsilon \in H^1(\Omega)$ with

$$-\nabla \cdot (A_\varepsilon(x)\nabla u_\varepsilon(x)) = x_1 - \frac{1}{2} \quad \text{in } \Omega$$
$$u_\varepsilon(x) = 0 \quad \text{on } \partial\Omega.$$

The scalar diffusion term A_ε is shown in Fig. 2. It is given by

$$A_\varepsilon(x) := (h \circ c_\varepsilon)(x) \quad \text{with } h(t) := \begin{cases} t^4 & \text{for } \frac{1}{2} < t < 1 \\ t^{\frac{3}{2}} & \text{for } 1 < t < \frac{3}{2} \\ t & \text{else} \end{cases} \tag{19}$$

and where

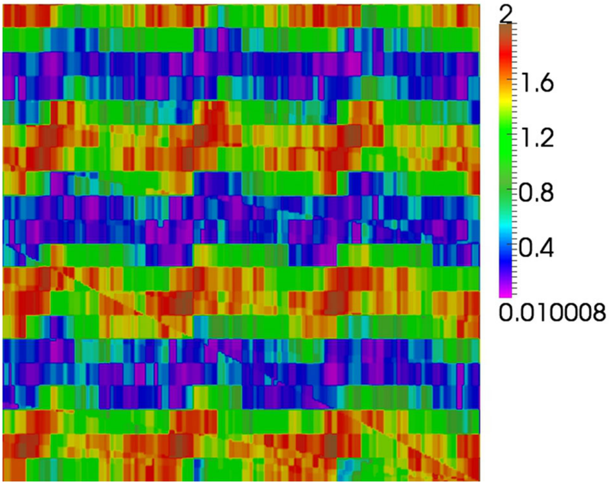

Fig. 2 Sketch of heterogeneous diffusion coefficient A_ε defined according to Eq. (19)

Table 1 Results for the errors between LOD approximations and reference solutions

H	k	$\|e_H\|^{\text{rel}}_{L^2(\Omega)}$	$\|e_h\|^{\text{rel}}_{L^2(\Omega)}$	$\|e_h\|^{\text{rel}}_{H^1(\Omega)}$	$\left\|e_H^{\text{PG}}\right\|^{\text{rel}}_{L^2(\Omega)}$	$\left\|e_h^{\text{PG}}\right\|^{\text{rel}}_{L^2(\Omega)}$	$\left\|e_h^{\text{PG}}\right\|^{\text{rel}}_{H^1(\Omega)}$
2^{-2}	0	0.3794	0.3772	0.6377	0.3778	0.3755	0.6375
2^{-2}	1/2	0.2756	0.2381	0.5312	0.2588	0.2269	0.5628
2^{-2}	1	0.2523	0.1445	0.3637	0.2544	0.1504	0.3642
2^{-2}	3/2	0.2514	0.1355	0.3125	0.2518	0.1380	0.3162
2^{-3}	0	0.2039	0.2037	0.5048	0.2037	0.2036	0.5048
2^{-3}	1	0.1100	0.0526	0.2278	0.1139	0.0619	0.2345
2^{-3}	2	0.1073	0.0423	0.1761	0.1078	0.0453	0.1807
2^{-3}	3	0.1070	0.0366	0.1567	0.1077	0.0399	0.1600
2^{-4}	0	0.0874	0.0873	0.3563	0.0874	0.0873	0.3563
2^{-4}	2	0.0353	0.0105	0.0932	0.0357	0.0123	0.0994
2^{-4}	4	0.0351	0.0082	0.0653	0.0353	0.0093	0.0680
2^{-4}	6	0.0351	0.0080	0.0634	0.0353	0.0091	0.0662

We define $e_h := u_h - u^{\text{G-LOD}}$ and $e_h^{\text{PG}} := u_h - u^{\text{PG-LOD}}$. Accordingly we define the errors between the reference solution and the coarse parts of the LOD approximations by $e_H := u_h - P_{L^2}(u^{\text{G-LOD}})$ (for the symmetric case) and $e_H^{\text{PG}} := u_h - P_{L^2}(u^{\text{PG-LOD}})$ (for the Petrov–Galerkin case). The reference solution u_h was obtained on a fine grid of mesh size $h = 2^{-6} \approx 0.0157 < \varepsilon$ which just resolves the micro structure of the coefficient A_ε. The number of 'coarse grid layers' is denoted by k and determines the patch size $U_k(T)$

$$c_\varepsilon(x_1, x_2) := 1 + \frac{1}{10}\sum_{j=0}^{4}\sum_{i=0}^{j}\left(\frac{2}{j+1}\cos\left(\left\lfloor i x_2 - \tfrac{x_1}{1+i}\right\rfloor + \left\lfloor \tfrac{i x_1}{\varepsilon}\right\rfloor + \left\lfloor \tfrac{x_2}{\varepsilon}\right\rfloor\right)\right).$$

The goal of the experiments is to investigate the accuracy of the PG-LOD, compared to the classical symmetric LOD. Moreover, we investigate the accuracy of the coarse part of the LOD approximation in terms of L^2-approximation properties (see Sect. 3.3 for a corresponding discussion).

In Table 1 we can see the results for a fine grid $\mathcal{T}_h$ with resolution $h = 2^{-6} < \varepsilon$ which just resolves the micro structure of the coefficient A_ε. Comparing the relative L^2- and H^1-errors for the G-LOD and the PG-LOD respectively (with the reference solution u_h), we observe that the errors are of similar size in each case. In general, we obtain slightly worse results for the PG-LOD, however the difference is so small that is does not justify the usage of the more memory-demanding (and more expensive) symmetric LOD. For both methods we observe the same nice error decay (in terms of the patch size) that was already predicted by the theoretical results. Comparing the relative L^2-errors between u_h and the coarse parts of the LOD-approximations, we observe that they already yield very good approximations. We also observe that they seem to be much more dominated by H-error contribution than by the θ^k-error contribution (i.e. the error coming from the decay). Using patches consisting of more than 8 fine element layers did not lead to any significant improvement, while there were still clear improvements visible for the other errors for the full G-LOD approximations.

Table 2 Results for the errors between LOD approximations and reference solutions

H	k	$\|e_H\|^{\text{rel}}_{L^2(\Omega)}$	$\|e_h\|^{\text{rel}}_{L^2(\Omega)}$	$\|e_h\|^{\text{rel}}_{H^1(\Omega)}$	$\left\|e_H^{\text{PG}}\right\|^{\text{rel}}_{L^2(\Omega)}$	$\left\|e_h^{\text{PG}}\right\|^{\text{rel}}_{L^2(\Omega)}$	$\left\|e_h^{\text{PG}}\right\|^{\text{rel}}_{H^1(\Omega)}$
2^{-2}	0	0.3840	0.3815	0.6434	0.3820	0.3796	0.6432
2^{-2}	1/8	0.2985	0.2781	0.5486	0.2957	0.2753	0.5513
2^{-2}	1/4	0.2852	0.2592	0.5578	0.2718	0.2472	0.5774
2^{-2}	1/2	0.2769	0.2392	0.5386	0.2607	0.2291	0.5722
2^{-2}	3/4	0.2676	0.2052	0.4784	0.2577	0.1972	0.4956
2^{-3}	0	0.2106	0.2103	0.5190	0.2103	0.2100	0.5190
2^{-3}	1/4	0.1480	0.1375	0.4510	0.1569	0.1469	0.4486
2^{-3}	1/2	0.1372	0.1163	0.3957	0.1305	0.1089	0.4029
2^{-3}	1	0.1138	0.0535	0.2308	0.1176	0.0628	0.2372
2^{-3}	3/2	0.1117	0.0399	0.1710	0.1126	0.0437	0.1761
2^{-4}	0	0.0988	0.0984	0.3854	0.0987	0.0983	0.3854
2^{-4}	1/2	0.0637	0.0592	0.2896	0.0500	0.0442	0.2934
2^{-4}	1	0.0406	0.0211	0.1613	0.0431	0.0263	0.1690
2^{-4}	2	0.0381	0.0109	0.0957	0.0385	0.0130	0.1017
2^{-4}	3	0.0380	0.0087	0.0726	0.0382	0.0099	0.0753

The errors are defined as in Table 1. The reference solution u_h was obtained on a fine grid of mesh size $h = 2^{-8} \approx 0.0039 \ll \varepsilon$ which fully resolves the micro structure of the coefficient A_ε. Again, the number of 'coarse grid layers' is denoted by k and determines the patch size $U_k(T)$

Furthermore, the linear convergence in H is clearly visible for $\|e_H\|^{\text{rel}}_{L^2(\Omega)}$ (respectively $\|e_H^{\text{PG}}\|^{\text{rel}}_{L^2(\Omega)}$) showing that the obtained error estimates seem to be indeed optimal.

The same observations can be made for the errors depicted in Table 2 for a fine grid $\mathcal{T}_h$ with resolution $h = 2^{-8} \ll \varepsilon$. Again, the results for the (symmetric) G-LOD are slightly better than the ones for the PG-LOD, but always of the same order. The exponential convergence in k for both realization is visualized in Fig. 3. It is clearly observable that there is no argument for using the G-LOD when dealing with patch communication issues which are storage demanding.

These findings are confirmed in the Figs. 4 and 5. In Fig. 4 we can see a visual comparison of the reference solution with the corresponding full LOD approximations (symmetric and Petrov–Galerkin). Both are almost not distinguishable for the investigated setting with $(h, H, k) = (2^{-8}, 2^{-4}, 2)$. Also the coarse parts of the LOD approximations already capture all the essential behavior of the reference solution. In Fig. 5 this is emphasized. Here, we compare the isolines between the reference solution and PG-LOD approximation (respectively its coarse part) and we observe that they are highly matching.

4.2 PG DG-LOD for the Buckley–Leverett equation

In this subsection we present the results of a two-phase flow simulation, based on solving the Buckley–Leverett equation as discussed in Sect. 3.6. Recall that, the Buckley–

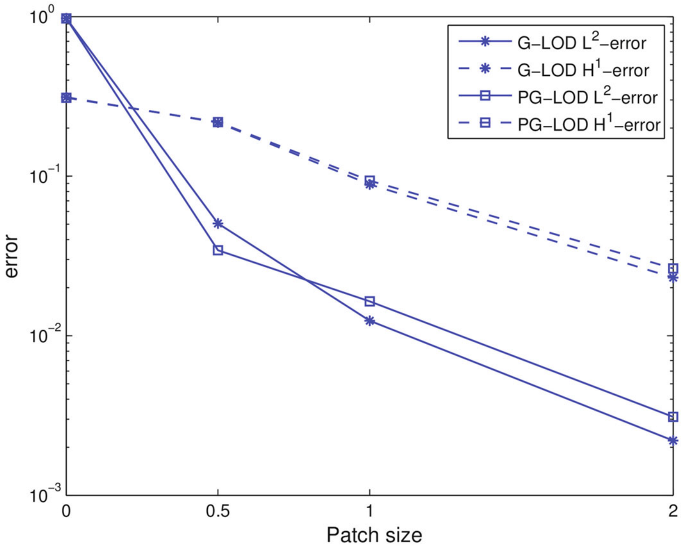

Fig. 3 The graphic visualizes the error decay in k. The results correspond to the results of Table 2 for $(h, H) = (2^{-8}, 2^{-4})$. We include $\|e_h\|^{\text{rel}}_{L^2(\Omega)}$, $\|e_h\|^{\text{rel}}_{H^1(\Omega)}$, $\|e_h^{\text{PG}}\|^{\text{rel}}_{L^2(\Omega)}$ and $\|e_h^{\text{PG}}\|^{\text{rel}}_{H^1(\Omega)}$. The x-axis depicts the localization parameter k and the y-axis the error "$\|e(k)\| - \|e(3)\|$" on the log-scale, where $\|e(k)\|$ denotes an error for k-layers (the error $\|e(3)\|$ is hence the limit reference)

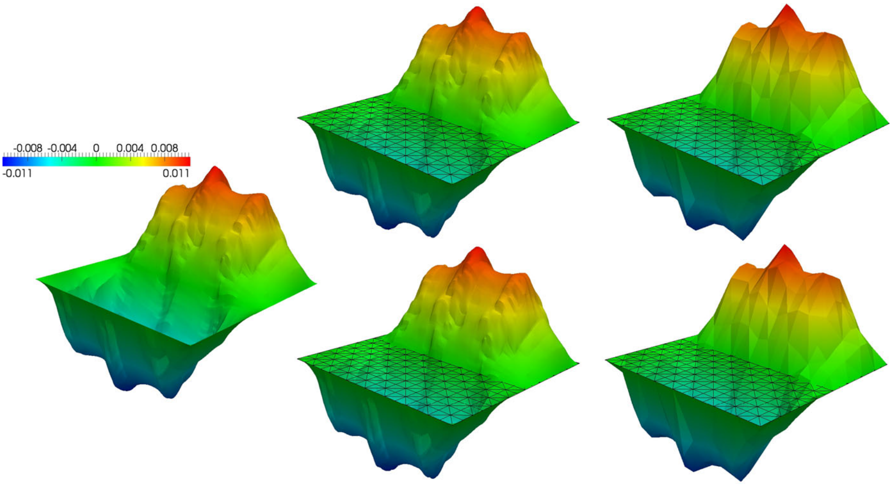

Fig. 4 The *left picture* shows the finite element reference solution u_h for $h = 2^{-8}$. The *remaining pictures* show LOD approximations for the case $(H, k) = (2^{-4}, 2)$, where k denotes the (*broken*) number of coarse layers. The two *top row pictures* show the full G-LOD approximation $u^{\text{G-LOD}}$ (*left*) and the coarse part of it, i.e. $P_{L^2}(u^{\text{G-LOD}})$ (*right*). The *bottom row* shows the full Petrov–Galerkin LOD approximation $u^{\text{PG-LOD}}$ (*left*) and the corresponding coarse part, i.e. $P_{L^2}(u^{\text{PG-LOD}})$ (*right*). The grid that is added to each of the pictures shows the coarse grid $\mathcal{T}_H$

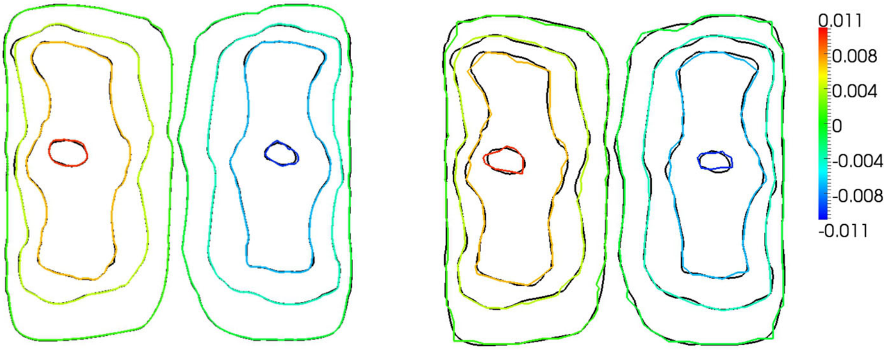

Fig. 5 The *pictures* depict a comparison of isolines. The *black lines* belong to the reference solution u_h for $h = 2^{-8}$. The *colored isolines in the left picture* belong to the PG-LOD approximation $u^{\text{PG-LOD}}$ and match almost perfectly with the one from the reference solution. The *right picture* shows the coarse part of $u^{\text{PG-LOD}}$, i.e. $P_{L^2}(u^{\text{PG-LOD}})$. We observe that the isolines still match nicely (color figure online)

Leverett equation has two parts, a hyperbolic equation for the saturation and a elliptic equation for the pressure. For that reason, we use the operator splitting technique IMPES, that we stated in Sect. 3.6. The elliptic pressure equation is solved by the PG DG-LOD for which a discontinuous linear finite element method is utilized that allows for recovering an elemental locally conservative normal flux. We emphasize that having a locally conservative flux is typically central for numerical schemes for solving hyperbolic partial differential equations. In this experiment we use an upwinding scheme.

Employing PG DG-LOD in this simulation proves to be a very efficient since the local correctors for the generalized basis functions only have to be computed once in a preprocessing step, this follows from the fact the saturation only influence the permeability on the macroscopic scale. The time stepping in the IMPES scheme using the PG DG-LOD for the is realized through Algorithm 2 below.

Set the end time T_{end}, number of update of the pressure n, number of explicit updates on each implicit step update m.

Algorithm 2: solveBuckleyLeverett($\mathcal{T}_H$, $\mathcal{T}_h$, T_{end}, n, m)

```
Set the initial values: $S = S_0$ and $i = 1$
Preprocessing step: Compute local corrections $Q_h^T$ for all $T \in \mathcal{T}_H$ with $\lambda(S) = 1$
**while** $t \leq T_{end}$ **do**
    Compute pressure $p$ using PG DG-LOD at $(t + T_{\text{end}}/(n))$
    Extract conservative flux $\mathbf{v}$
    **while** $t \leq iT_{end}/n$ **do**
        Compute saturation $S$ at $(t + T_{\text{end}}/(nm))$
        Update time: $t + T_{\text{end}}/(nm) \mapsto t$
    **end**
    $i + 1 \mapsto i$
**end**
```

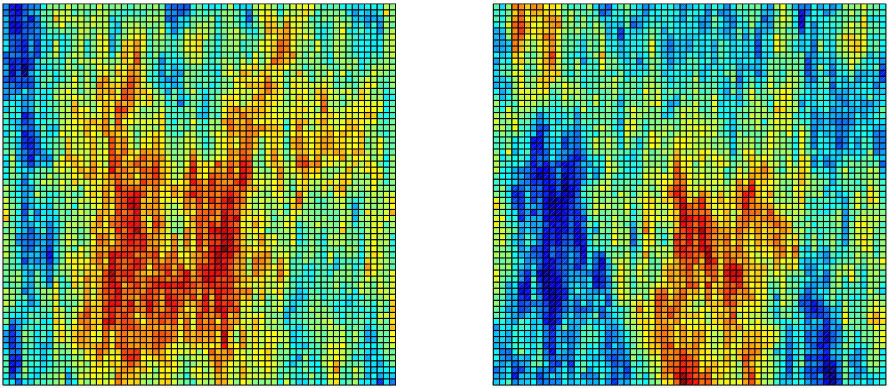

Fig. 6 The permeability structure of K_i in log scale with, $\beta_0/\alpha_0 \approx 5 \times 10^5$ for $i = 1$ (*left*) and $\beta_0/\alpha_0 \approx 4 \times 10^5$ for $i = 2$ (*right*)

Table 3 The resulting error in relative L^2-norm between S and S^{ref}, where S is obtained using PG DG-LOD for the pressure computed on $\mathcal{T}_H$ and S^{ref} is the reference solution computed on $\mathcal{T}_h$

Data	$\|e(T_1)\|_{L^2(\Omega)}$	$\|e(T_2)\|_{L^2(\Omega)}$	$\|e(T_3)\|_{L^2(\Omega)}$
K_1	0.088	0.073	0.070
K_2	0.058	0.087	0.079

We have $T_1 = 0.05$, $T_2 = 0.25$ and $T_3 = 0.45$

In the numerical experiment we consider the domain Ω to be the unit square. The permeability K_i for $i = 1, 2$ is given by layer 21 and 31 of the Society of Petroleum Engineering comparative permeability data (available on http://www.spe.org/web/csp), projected on a uniform mesh with resolution 2^{-6} as illustrated in Fig. 6.

We consider a microscopic partition $\mathcal{T}_h$ with mesh size size $h = 2^{-8}$ and a macroscopic partition $\mathcal{T}_H$ with mesh size $H = 2^{-i}$ for $i = 3, 4, 5, 6$. The patch size is chosen such that the overall H convergence for the PG DG-LOD is not effected. A reference solution to the Buckley–Leverett equation is obtained when both the pressure and saturation equation are computed on $\mathcal{T}_h$, compared to using Algorithm 2 where both the pressure and saturation equation are computed on $\mathcal{T}_H$. We consider the following setup. For the pressure equation we use the boundary condition $p = 1$ for the left boundary, $p = 0$ for the right boundary, $K\lambda(S)\nabla p = 0$ otherwise, and the source terms $q_w = q_n = 0$. For the saturation the initial value is $S = 1$ on the left boundary and 0 elsewhere. The error is defined by $e(\cdot, t) := S(\cdot, t) - S^{\text{rel}}(\cdot, t)$, where $S(\cdot, t)$ is the solution obtained by Algorithm 2 (at time t) and $S^{\text{rel}}(\cdot, t)$ is the reference solution (at time t). The errors are measured in the L^2-norm. In Table 3 we fix the coarse mesh size to be $H = 2^{-5}$, and compute the error for the permeabilities K_1 and K_2 at the times $T_1 := 0.05$, $T_2 := 0.25$ and $T_3 := 0.45$. A graphical comparison is shown in Figs. 7 and 8. The errors in the L^2-norm is less than 0.1 for both permeabilities at all times which is quite remarkable since the coarse mesh $\mathcal{T}_H$ for $H = 2^{-5}$ does not resolve the data. In Table 4 we consider the test case involving permeability K_1. We

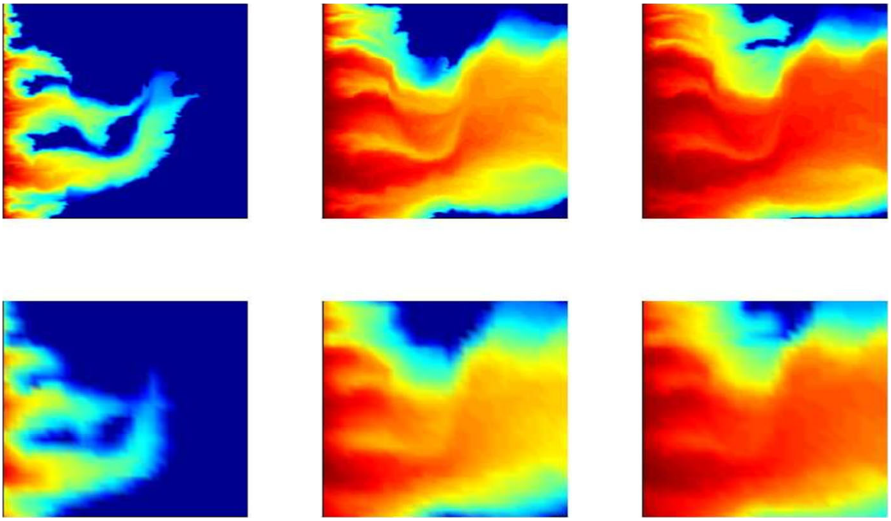

Fig. 7 The saturation profile using PG DG-LOD for the pressure equation on the grid $\mathcal{T}_H$ (*bottom*) and the reference solution on the grid $\mathcal{T}_h$ (*upper*) at time $T_1 = 0.05$ (*left*), $T_2 = 0.25$ (*middle*), and $T_3 = 0.45$ (*right*) using permeability K_1

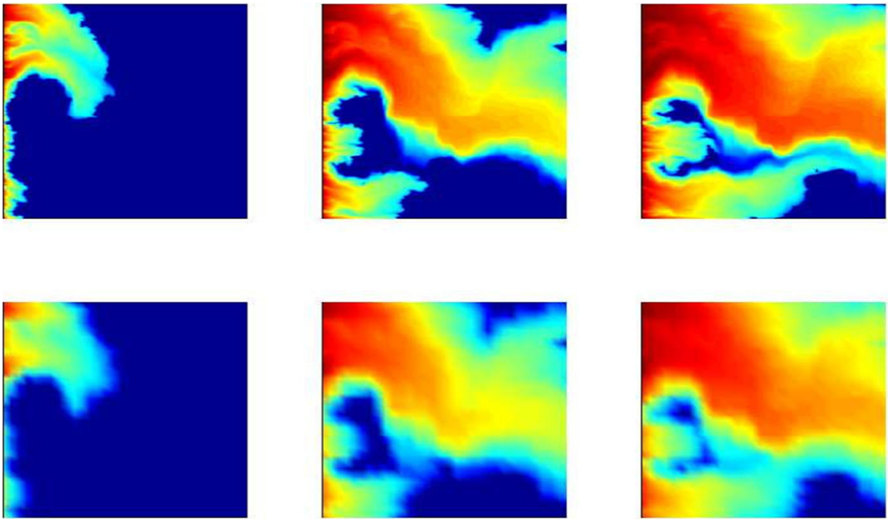

Fig. 8 The saturation profile using PG DG-LOD for the pressure equation on the grid $\mathcal{T}_H$ (*bottom*) and the reference solution on the grid $\mathcal{T}_h$ (*upper*) at time $T_1 = 0.05$ (*left*), $T_2 = 0.25$ (*middle*), and $T_3 = 0.45$ (*right*) using permeability K_2

present the L^2-errors at $t = T_2$ for different values of H. We basically observe a linear convergence rate in H/h (for fixed h) which is just what we would expect (since we only use the coarse part of the LOD pressure approximation).

Table 4 We consider the test case involving K_1

H	$\|e(T_2)\|_{L^2(\Omega)}$
2^{-3}	0.220
2^{-4}	0.113
2^{-5}	0.073
2^{-6}	0.048

The table depicts relative L^2-errors between S and S^{ref} at $T_2 = 0.25$ for different values of the coarse mesh size H. Here, S^{ref} denotes the reference solution computed on T_h with $h = 2^{-8}$ and S denotes the numerical approximation obtained with the IMPES scheme, using the PG DG-LOD for solving the pressure equation (with coarse mesh $\mathcal{T}_H$). We pick $k = \lceil 2|\log(H)|\rceil$

5 Proofs of the main results

In this proof section we will frequently exploit the estimate

$$\|v_h\|_{L^2(\Omega)} \lesssim |||v_h|||_h \quad \text{for all } v_h \in V_h, \tag{20}$$

which is a conclusion from assumption (A7). Let $I_H^{-1} := (I_H|_{V_H})^{-1}$, then (20) can be verified as follows by using (A7).

$$\begin{aligned}\|v_h\|_{L^2(\Omega)} &\leq \|v_h - I_H(v_h)\|_{L^2(\Omega)} + \|I_H(v_h)\|_{L^2(\Omega)}\\ &\lesssim H|||v_h|||_h + \|(I_H \circ I_H^{-1} \circ I_H)(v_h)\|_{L^2(\Omega)}\\ &\lesssim H|||v_h|||_h + |||(I_H^{-1} \circ I_H)(v_h)|||_H \lesssim H|||v_h|||_h + |||I_H(v_h)|||_H\\ &\lesssim H|||v_h|||_h + |||v_h|||_h.\end{aligned}$$

5.1 Proof of Theorem 1

The arguments for establishing the error estimate in $|||\cdot|||_h$-norm is analogous to the standard case, see for example [35] or [19]. We only recall the main arguments.

Proof (Proof of Theorem 1) Let $u_H^{\text{G-LOD}} = (u_H + Q_h(u_H)) \in V^{\text{ms}}$ be the Galerkin LOD solution governed by (10). Utilizing the notation in (A8), we set $u_{H,\Omega} \in V_H$ to satisfy

$$a_h(u_{H,\Omega} + Q_h^{\Omega}(u_{H,\Omega}), \Phi_H + Q_h^{\Omega}(\Phi_H)) = (f, \Phi_H + Q_h^{\Omega}(\Phi_H)) \quad \text{for all } \Phi_H \in V_H$$

and define $e_h := u_{H,\Omega} + Q_h^{\Omega}(u_{H,\Omega}) - u_h$. Using Galerkin orthogonality, we obtain $a_h(e_h, \Phi) = 0$ for all $\Phi \in V_{\Omega}^{\text{ms}}$ and hence $e_h \in W_h$ (i.e. $I_H(e_h) = 0$). This implies $|||e_h|||_h^2 \lesssim a_h(e_h, e_h) = (f, e_h) = (f, e_h - I_H(e_h)) \lesssim H\|f\|_{L^2(\Omega)}\, |||e_h|||_h$ and consequently by energy minimization

$$\begin{aligned}|||u_H^{\text{G-LOD}} - u_h|||_h &= |||u_H + Q_h(u_H) - u_h|||_h \lesssim |||u_{H,\Omega} + Q_h(u_{H,\Omega}) - u_h|||_h \\ &\leq |||e_h|||_h + \left|\left|\left| Q_h^{\Omega}(u_{H,\Omega}) - Q_h(u_{H,\Omega}) \right|\right|\right|_h \\ &\overset{\text{(A8)}}{\lesssim} H\|f\|_{L^2(\Omega)} + (1/H)^p k^{d/2}\theta^k \left|\left|\left| u_{H,\Omega} + Q_h^{\Omega}(u_{H,\Omega}) \right|\right|\right|_h .\end{aligned}$$

The bound $|||u_{H,\Omega} + Q_h^{\Omega}(u_{H,\Omega})|||_h \lesssim \|f\|_{L^2(\Omega)}$ finishes the energy-norm estimate. The estimate in the L^2-norm is established in a similar fashion using (20). □

5.2 Proof of Theorem 2

We begin with stating and proving a lemma that is required to establish the a priori error estimate.

Lemma 3 *For all $v^{\text{ms}} \in V_{\Omega}^{\text{ms}}$ with $v^{\text{ms}} = v_H + v^f$, where $v_H \in V_H$ and $v^f \in W_h$, we have*

$$\|v^f\|_{L^2(\Omega)} \lesssim H|||v^{\text{ms}}|||_h. \tag{21}$$

Proof Because of $I_H(v^f) = 0$ and $(I_H^{-1} \circ I_H)(v_H) = v_H$,

$$v^f = v^f - I_H(v^f) + v_H - (I_H^{-1} \circ I_H)(v_H + v^f) + I_H(v_H + v^f) - I_H(v_H),$$

and therefore with $I_H = I_H \circ I_H^{-1} \circ I_H$ and (A7),

$$\begin{aligned}\|v^f\|_{L^2(\Omega)} &\leq \|v^{\text{ms}} - I_H(v^{\text{ms}})\|_{L^2(\Omega)} + \left\| \left(I_H^{-1} \circ I_H\right)(v^{\text{ms}}) - I_H(v^{\text{ms}}) \right\|_{L^2(\Omega)} \\ &\lesssim H|||v^{\text{ms}}|||_h + \left\| \left(I_H^{-1} \circ I_H\right)(v^{\text{ms}}) - \left(I_H \circ I_H^{-1} \circ I_H\right)(v^{\text{ms}}) \right\|_{L^2(\Omega)} \\ &\lesssim H|||v^{\text{ms}}|||_h + H \left|\left|\left| \left(I_H^{-1} \circ I_H\right)(v^{\text{ms}}) \right|\right|\right|_H \\ &\lesssim H|||v^{\text{ms}}|||_h.\end{aligned}$$

In the last step we used again the stability estimates for I_H^{-1} and I_H in (A7). □

Proof (Proof of Theorem 2) Let $u_{H,\Omega}^{\text{G-LOD}}$ and $u_{H,\Omega}^{\text{PG-LOD}}$ be respectively the solution of (10) and (12) for $U(T) = \Omega$. As in the statement of the theorem, $u_H^{\text{PG-LOD}}$ is the solution of (12) for $U(T) = U_k(T)$. By adding and subtracting appropriate terms and applying triangle inequality, we arrive at

$$\left|\left|\left| u_h - u_H^{\text{PG-LOD}} \right|\right|\right|_h \leq \text{I}_1 + \text{I}_2 + \text{I}_3,$$

where we set $\text{I}_1 = |||u_h - u_{H,\Omega}^{\text{G-LOD}}|||_h$, $\text{I}_2 = |||u_{H,\Omega}^{\text{G-LOD}} - u_{H,\Omega}^{\text{PG-LOD}}|||_h$, and $\text{I}_3 = |||u_{H,\Omega}^{\text{PG-LOD}} - u_H^{\text{PG-LOD}}|||_h$. In the following, we estimate these three terms. Because

$e^{(1)} := (u_h - u_{H,\Omega}^{\text{G-LOD}}) \in W_h$ (cf. proof of Theorem 1) and by applying the Galerkin orthogonality, we get

$$\begin{aligned}\mathrm{I}_1^2 &\lesssim a_h(e^{(1)}, e^{(1)}) = a_h(u_h, e^{(1)}) \\ &\quad = (f, e^{(1)} - I_H(e^{(1)})) \lesssim H\|f\|_{L^2(\Omega)} \quad |||e^{(1)}|||_h \leq H\|f\|_{L^2(\Omega)} \quad \mathrm{I}_1,\end{aligned} \tag{22}$$

i.e. $\mathrm{I}_1 \lesssim H\|f\|$. Furthermore, $e^{(2)} := (u_{H,\Omega}^{\text{PG-LOD}} - u_{H,\Omega}^{\text{G-LOD}}) \in V_\Omega^{\text{ms}}$ and the splitting $e^{(2)} = e_H^{(2)} + e_f^{(2)}$ with $e_H^{(2)} \in V_H$ and $e_f^{(2)} \in W_h$ (i.e. $I_H(e_f^{(2)}) = 0$) holds true. Because $a_h(u_{H,\Omega}^{\text{PG-LOD}}, e_f^{(2)}) = 0$, we obtain

$$\begin{aligned}\mathrm{I}_2^2 &\lesssim a_h(e^{(2)}, e^{(2)}) \\ &= a_h\left(u_{H,\Omega}^{\text{PG-LOD}}, e_H^{(2)}\right) - a_h\left(u_{H,\Omega}^{\text{G-LOD}}, e^{(2)}\right) = \left(f, e_H^{(2)} - e^{(2)}\right) = -\left(f, e_f^{(2)}\right),\end{aligned} \tag{23}$$

where $(f, e_f^{(2)}) \leq \|f\|_{L^2(\Omega)} \|e_f^{(2)}\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)} H|||e^{(2)}|||_h = H\|f\|_{L^2(\Omega)} \mathrm{I}_2$ by Lemma 3. Again, we conclude that $\mathrm{I}_2 \lesssim H\|f\|_{L^2(\Omega)}$. It remains to estimate I_3 for which we define $e^{(3)} := u_{H,\Omega}^{\text{PG-LOD}} - u_H^{\text{PG-LOD}}$. To simplify the notation, we subsequently denote (according to the definitions of V^{ms} and V_Ω^{ms})

$$u_H^{\text{PG-LOD}} = u_H + Q_h(u_H) \quad \text{and} \quad u_{H,\Omega}^{\text{PG-LOD}} = u_H^\Omega + Q_h^\Omega\left(u_H^\Omega\right),$$

where $u_H \in V_H$ and $u_H^\Omega \in V_H$. By the definition of problem (12) we have

$$a_h\left(u_H^{\text{PG-LOD}}, \Phi_H\right) = (f, \Phi_H) = a_h\left(u_{H,\Omega}^{\text{PG-LOD}}, \Phi_H\right). \tag{24}$$

On the other hand, by the definition of $Q_h^\Omega = -P_h$ (see Remark 1) and since $Q_h(\Phi_H) \in W_h$ we get

$$a_h\left(u_{H,\Omega}^{\text{PG-LOD}}, Q_h(\Phi_H)\right) = 0. \tag{25}$$

Combining (24) and (25) we get the equality

$$\begin{aligned}a_h\left(u_H^{\text{PG-LOD}}, \Phi_H + Q_h(\Phi_H)\right) &= a_h\left(u_H^{\text{PG-LOD}}, Q_h(\Phi_H)\right) \\ &\quad + a_h\left(u_{H,\Omega}^{\text{PG-LOD}}, \Phi_H + Q_h(\Phi_H)\right).\end{aligned}$$

We use this equality cast u_H as a unique solution of a self-adjoint variational equation expressed as

$$a_h(u_H + Q_h(u_H), \Phi_H + Q_h(\Phi_H)) = F_{u_H, u_H^\Omega}(\Phi_H) \quad \text{for all } \Phi_H \in V_H,$$

where F_{u_H, u_H^Ω} is a given fixed data function written as

$$F_{u_H, u_H^\Omega}(\Phi_H) = a_h(u_H + Q_h(u_H), Q_h(\Phi_H)) + a_h\left(u_H^\Omega + Q_h^\Omega\left(u_H^\Omega\right), \Phi_H + Q_h(\Phi_H)\right).$$

Since this problem is self-adjoint, we get that u_H is equally the minimizer in V_H of the functional

$$\begin{aligned} J(\Phi_H) := a_h &\left(\Phi_H + Q_h(\Phi_H) - u_H^\Omega - Q_h^\Omega\left(u_H^\Omega\right), \Phi_H + Q_h(\Phi_H) - u_H^\Omega - Q_h^\Omega\left(u_H^\Omega\right)\right) \\ &-2a_h(u_H + Q_h(u_H), Q_h(\Phi_H)). \end{aligned}$$

Hence we obtain

$$\begin{aligned} \alpha \mathrm{I}_3^2 &= \alpha|||e^{(3)}|||_h^2 \\ &\leq a_h(e^{(3)}, e^{(3)}) \\ &= J(u_H) + 2a_h(u_H + Q_h(u_H), Q_h(u_H)) \\ &\leq J\left(u_H^\Omega\right) + 2a_h(u_H + Q_h(u_H), Q_h(u_H)) \\ &= a_h\left(Q_h\left(u_H^\Omega\right) - Q_h^\Omega\left(u_H^\Omega\right), Q_h\left(u_H^\Omega\right) - Q_h^\Omega\left(u_H^\Omega\right)\right) \\ &\quad - 2a_h\left(u_H + Q_h(u_H), Q_h(u_H) - Q_h\left(u_H^\Omega\right)\right) \\ &= \mathrm{I}_{31} + \mathrm{I}_{32}, \end{aligned} \tag{26}$$

where

$$\begin{aligned} \mathrm{I}_{31} &= a_h\left(Q_h\left(u_H^\Omega\right) - Q_h^\Omega\left(u_H^\Omega\right), Q_h\left(u_H^\Omega\right) - Q_h^\Omega\left(u_H^\Omega\right)\right) \\ \mathrm{I}_{32} &= a_h\left(Q_h(u_H) - Q_h^\Omega(u_H), Q_h(u_H) - Q_h\left(u_H^\Omega\right)\right). \end{aligned}$$

By the boundedness of $a_h(\cdot,\cdot)$ and applying (9) we get

$$\mathrm{I}_{31} \lesssim \left|\left|\left|Q_h\left(u_H^\Omega\right) - Q_h^\Omega\left(u_H^\Omega\right)\right|\right|\right|_h^2 \lesssim k^p\theta^{2k}(1/H)^{2p}\left|\left|\left|u_H^\Omega + Q_h^\Omega\left(u_H^\Omega\right)\right|\right|\right|_h^2. \tag{27}$$

We now need to estimate $u_{H,\Omega}^{\text{PG-LOD}} = u_H^\Omega + Q_h^\Omega(u_H^\Omega)$. By the inf-sup condition and Lemma 3,

$$\begin{aligned} \left|\left|\left|u_{H,\Omega}^{\text{PG-LOD}}\right|\right|\right|_h^2 &\lesssim a_h\left(u_{H,\Omega}^{\text{PG-LOD}}, u_{H,\Omega}^{\text{PG-LOD}}\right) \\ &= a\left(u_{H,\Omega}^{\text{PG-LOD}}, u_H^\Omega\right) \\ &= \left(f, u_H^\Omega\right) \\ &= \left(f, u_{H,\Omega}^{\text{PG-LOD}}\right) - \left(f, Q_h^\Omega(u_H^\Omega)\right) \\ &\lesssim (1+H)\|f\|_{L^2(\Omega)}\left|\left|\left|u_{H,\Omega}^{\text{PG-LOD}}\right|\right|\right|_h, \end{aligned} \tag{28}$$

and thus combining it with (27) yields

$$\mathrm{I}_{31} \lesssim k^d \theta^{2k} (1/H)^{2p} \|f\|^2_{L^2(\Omega)} \tag{29}$$

Furthermore, in a similar fashion we use the boundedness of $a_h(\cdot,\cdot)$ and (9) to get

$$\begin{aligned} \mathrm{I}_{32} &\lesssim |||Q_h(u_H) - Q_h^\Omega(u_H)|||_h \;\; |||Q_h(u_H) - Q_h\left(u_H^\Omega\right)|||_h \\ &\lesssim k^{d/2}\theta^k (1/H)^p \left|\left|\left|u_H^{\text{PG-LOD}}\right|\right|\right|_h \;\; |||Q_h(u_H) - Q_h\left(u_H^\Omega\right)|||_h \end{aligned} \tag{30}$$

By adding and subtracting appropriate terms and applying triangle inequality

$$\begin{aligned} &|||Q_h(u_H) - Q_h\left(u_H^\Omega\right)|||_h \\ &\quad\le |||\left(Q_h - Q_h^\Omega\right)(u_H)|||_h + |||Q_h^\Omega\left(u_H - u_H^\Omega\right)|||_h + |||\left(Q_h^\Omega - Q_h\right)\left(u_H^\Omega\right)|||_h . \end{aligned} \tag{31}$$

We use (9) to estimate the first and last terms in (31) to yield

$$\begin{aligned} &|||\left(Q_h - Q_h^\Omega\right)(u_H)|||_h + |||\left(Q_h^\Omega - Q_h\right)\left(u_H^\Omega\right)|||_h \\ &\quad\lesssim k^{d/2}\theta^k (1/H)^p \left(\left|\left|\left|u_H^{\text{PG-LOD}}\right|\right|\right|_h + \left|\left|\left|u_{H,\Omega}^{\text{PG-LOD}}\right|\right|\right|_h\right). \end{aligned} \tag{32}$$

Moreover, by the $|||\cdot|||_h$-stability of Q_h^Ω [which holds true since $Q_h^\Omega = -P_h$ with P_h being the orthogonal projection defined in (4)], we have

$$\begin{aligned} |||Q_h^\Omega\left(u_H - u_H^\Omega\right)|||_h &\lesssim |||u_H - u_H^\Omega|||_h \\ &= \left|\left|\left|\left((I_H|_{V_H})^{-1} \circ I_H\right)(e^{(3)})\right|\right|\right|_h \lesssim C_{H,h}|||e^{(3)}|||_h. \end{aligned} \tag{33}$$

Putting back (33) and (32) to (31) and place it in (30) gives

$$\begin{aligned} \mathrm{I}_{32} &\lesssim k^d\theta^{2k}(1/H)^{2p} \left|\left|\left|u_H^{\text{PG-LOD}}\right|\right|\right|_h \left(\left|\left|\left|u_H^{\text{PG-LOD}}\right|\right|\right|_h + \left|\left|\left|u_{H,\Omega}^{\text{PG-LOD}}\right|\right|\right|_h\right) \\ &\quad + k^{d/2}\theta^k(1/H)^p \left|\left|\left|u_H^{\text{PG-LOD}}\right|\right|\right|_h \;\; C_{H,h}|||e^{(3)}|||_h \\ &\lesssim k^d\theta^{2k}(1/H)^{2p} \left(\left|\left|\left|u_H^{\text{PG-LOD}}\right|\right|\right|_h^2 + \left|\left|\left|u_{H,\Omega}^{\text{PG-LOD}}\right|\right|\right|_h^2\right) \\ &\quad + \frac{C_{H,h}^2}{\delta} k^d\theta^{2k}(1/H)^{2p} \left|\left|\left|u_H^{\text{PG-LOD}}\right|\right|\right|_h^2 + \frac{\delta}{4}|||e^{(3)}|||_h^2, \end{aligned} \tag{34}$$

where in the last step we use the Young's inequality for both terms, and in particular for the second term, inserting a sufficiently small $\delta > 0$ so that we can later on hide the term $\frac{\delta}{4}|||e^{(3)}|||_h^2$ in the left hand side of (26). Note that the choice of δ is independent of H, h or k. Rearranging and collecting common terms in the last inequality gives

$$\mathrm{I}_{32} \lesssim k^d \theta^{2k} (1/H)^{2p} \left(\left(1 + \frac{C_{H,h}^2}{\delta} \right) \left|\left|\left| u_H^{\text{PG-LOD}} \right|\right|\right|^2 + \left|\left|\left| u_{H,\Omega}^{\text{PG-LOD}} \right|\right|\right|^2 \right) + \frac{\delta}{4} |||e^{(3)}|||_h^2,$$

so that we need to estimate $|||u_H^{\text{PG-LOD}}|||_h$ and $|||u_{H,\Omega}^{\text{PG-LOD}}|||_h$, respectively. The stability of the second piece was established in (28), while the stability of the first piece is achieved by employing (A9) and (A7) in

$$\bar{\alpha} \left|\left|\left| u_H^{\text{PG-LOD}} \right|\right|\right|_h |||u_H|||_H \lesssim a_h \left(u_H^{\text{PG-LOD}}, u_H \right) = (f, u_H) \lesssim \|f\|_{L^2(\Omega)} |||u_H|||_H.$$

From which we conclude that

$$\mathrm{I}_{32} \lesssim k^d \theta^{2k} (1/H)^{2p} \left(\left(1 + \frac{C_{H,h}^2}{\delta} \right) (1 + \bar{\alpha}^{-1}) \|f\|^2 \right) + \frac{\delta}{4} \mathrm{I}_3^2.$$

To summarize, putting this last inequality and (29)–(26) and choosing sufficiently small δ gives

$$\mathrm{I}_3 \lesssim k^{d/2} \theta^k (1/H)^p \left(\left(1 + \frac{C_{H,h}}{\delta} \right) (1 + \bar{\alpha}^{-1}) \|f\| \right),$$

combining it with the existing estimates for I_1 and I_2 proves the error estimate in $|||\cdot|||_h$. Moreover, the estimate in the L^2-norm is established in a similar fashion. This completes the proof of the theorem. □

5.3 Proof of Lemmas 1 and 2

Next, we prove the inf-sup stability of the continuous Galerkin LOD in Petrov–Galerkin formulation.

Proof (Proof of Lemma 1) Let $\Phi^{\text{ms}} \in V^{\text{ms}}$ be an arbitrary element. To prove the inf-sup condition, we aim to show that

$$\frac{a_h(\Phi^{\text{ms}}, \Phi_H)}{|||\Phi_H|||_h} \geq \alpha(k) |||\Phi^{\text{ms}}|||_h \quad \text{for } \Phi_H = \left((I_H|_{V_H})^{-1} \circ I_H \right) (\Phi^{\text{ms}}). \qquad (35)$$

Let therefore $U(T) = U_k(T)$ for fixed $k \in \mathbb{N}$. By the definitions of V^{ms} and Φ_H, we have $\Phi^{\text{ms}} = \Phi_H + Q_h(\Phi_H)$, where $Q_h(\Phi_H)$ denotes the corresponding corrector given by (7). By $Q_h^{\Omega}(\Phi_H)$ we denote the corresponding global corrector for the case $U(T) = \Omega$ and the local correctors are denoted by $Q_h^{\Omega,T}(\Phi_H)$. First, we observe that by $|||\cdot|||_h = |||\cdot|||_H$

$$|||\Phi_H|||_h = \left|\left|\left| \left((I_H|_{V_H})^{-1} \circ I_H \right) (\Phi^{\text{ms}}) \right|\right|\right|_h \lesssim |||\Phi^{\text{ms}}|||_h, \qquad (36)$$

where we used the $|||\cdot|||_h$-stability of I_H and $(I_H|_{V_H})^{-1}$ according to (A7). Consequently, Eq. (36) implies

$$|||Q_h(\Phi_H)|||_h \leq |||\Phi^{\mathrm{ms}}|||_h + |||\Phi_H|||_h \lesssim |||\Phi^{\mathrm{ms}}|||_h, \tag{37}$$

and thus

$$\begin{aligned} a_h(\Phi^{\mathrm{ms}}, \Phi_H) &= a_h(\Phi^{\mathrm{ms}}, \Phi^{\mathrm{ms}}) - a_h\left(\Phi^{\mathrm{ms}}, Q_h(\Phi_H)\right) \\ &\geq \alpha |||\Phi^{\mathrm{ms}}|||_h^2 - a_h\left(\Phi^{\mathrm{ms}}, Q_h(\Phi_H)\right) \\ &\geq C\alpha |||\Phi_H|||_h\, |||\Phi^{\mathrm{ms}}|||_h - a_h\left(\Phi^{\mathrm{ms}}, Q_h(\Phi_H)\right), \end{aligned} \tag{38}$$

where we have used (36) again to bound $|||\Phi^{\mathrm{ms}}|||_h$ from below. Note here that C denotes a generic constant. It remains to bound $a_h(\Phi^{\mathrm{ms}}, Q_h(\Phi_H))$. By the orthogonality of V_Ω^{ms} and W_h we have

$$a_h\left(\Phi_H + Q_h^\Omega(\Phi_H), Q_h(\Phi_H)\right) = 0, \tag{39}$$

and since $a_h(\cdot,\cdot)$ is such that $a_h(v_h, w_h) = 0$ for all $v_h, w_h \in V_h$ with the property $\mathrm{supp}(v_h) \cap \mathrm{supp}(w_h) = \emptyset$ we get by the definition of $Q_h(\Phi_H)$ for every $w_h^T \in W_h(T)$

$$\begin{aligned} a_h\left(\Phi_H + Q_h(\Phi_H), w_h^T\right) &= \sum_{K\in\mathcal{T}_H} \left(a_h^K\left(\Phi_H, w_h^T\right) + a_h\left(Q_h(\Phi_H), w_h^T\right)\right) \\ &= \left(\sum_{K\in\mathcal{T}_H} a_h^K\left(\Phi_H, w_h^T\right)\right) + a_h\left(Q_h^T(\Phi_H), w_h^T\right) \\ &= a_h\left(\Phi_H + Q_h^T(\Phi_H), w_h^T\right) \\ &= 0. \end{aligned} \tag{40}$$

Using both equalities above and by the boundedness of $a_h(\cdot,\cdot)$ and applying (37) yields

$$\begin{aligned} &a_h\left(\Phi^{\mathrm{ms}}, Q_h(\Phi_H)\right) \\ &\quad = a_h\left(\Phi_H + Q_h^\Omega(\Phi_H), Q_h(\Phi_H)\right) + a_h\left(Q_h(\Phi_H) - Q_h^\Omega(\Phi_H), Q_h(\Phi_H)\right) \\ &\quad = a_h\left(Q_h(\Phi_H) - Q_h^\Omega(\Phi_H), Q_h(\Phi_H) - w_h\right) \\ &\quad \leq \left|\left|\left|Q_h(\Phi_H) - Q_h^\Omega(\Phi_H)\right|\right|\right|_h \frac{|||Q_h(\Phi_H) - w_h|||_h}{|||Q_h(\Phi_H)|||_h} |||\Phi^{\mathrm{ms}}|||_h. \end{aligned} \tag{41}$$

We next estimate $|||Q_h(\Phi_H) - Q_h^\Omega(\Phi_H)|||_h$ by applying (14) and establishing an analog of (15) for $Q_h^{\Omega,T}(\Phi_H)$ expressed as

$$\left|\left|\left|Q_h^{\Omega,T}(\Phi_H)\right|\right|\right|_h^2 \lesssim |||\Phi_H|||_{h,T} \left|\left|\left|Q_h^{\Omega,T}(\Phi_H)\right|\right|\right|_h, \tag{42}$$

giving (for $k > 0$)

$$\left|\left|\left|Q_h(\Phi_H) - Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h \lesssim k^{d/2}\theta^k \left(\sum_{T\in\mathcal{T}_H} \left|\left|\left|Q_h^{\Omega,T}(\Phi_H)\right|\right|\right|_h^2\right)^{1/2}$$
$$\lesssim k^{d/2}\theta^k \left(\sum_{T\in\mathcal{T}_H} |||\Phi_H|||_{h,T}^2\right)^{1/2}$$
$$\lesssim k^{d/2}\theta^k |||\Phi_H|||_h. \tag{43}$$

Thus we end up with

$$a_h(\Phi^{\mathrm{ms}}, Q_h(\Phi_H)) \lesssim \left(\frac{|||Q_h(\Phi_H) - w_h|||_h}{|||Q_h(\Phi_H)|||_h}\right) k^{d/2}\theta^k |||\Phi_H|||_h \; |||\Phi^{\mathrm{ms}}|||_h, \tag{44}$$

which when combined with (38) implies that there exist positive generic constants C_1, C_2 (independent of H and k) such that

$$\frac{a_h(\Phi^{\mathrm{ms}}, \Phi_H)}{|||\Phi_H|||_h \quad |||\Phi^{\mathrm{ms}}|||_h} \geq C_1\alpha - C_2 k^{d/2}\theta^k \inf_{w_h\in W_h^T} \frac{|||Q_h(\Phi_H) - w_h|||_h}{|||Q_h(\Phi_H)|||_h}. \tag{45}$$

Since $\inf_{w_h\in W_h^T} \frac{|||Q_h(\Phi_H)-w_h|||_h}{|||Q_h(\Phi_H)|||_h} = 0$ for $k = 0$, estimate (45) holds for all $k \in \mathbb{N}$ and the condition $k > 0$ is not required. The relation $Q_h(\Phi_H) = \Phi^{\mathrm{ms}} - ((I_H|_{V_H})^{-1} \circ I_H)(\Phi^{\mathrm{ms}})$ finishes the proof. □

Finally, we prove the inf-sup stability of the discontinuous Galerkin LOD in Petrov–Galerkin formulation.

Proof (Proof of Lemma 2) The main arguments are similar as in the proof of Lemma 1. Set $n := (m + 3)/2$. Let $\Phi^{\mathrm{ms}} = \Phi_H + Q_h(\Phi_H) \in V^{\mathrm{ms}}$ be an arbitrary element and let $U(T) = U_k(T)$ for fixed $k \gtrsim n|\log(H)|$. By the assumptions on $\mathcal{T}_H$ and $\mathcal{T}_h$ and by the definitions of $|||\cdot|||_h$ and $|||\cdot|||_H$ it is easy to see that

$$|||\Phi_H|||_h \lesssim H^{(1-m)/2}|||\Phi_H|||_H \quad \text{and} \quad |||\Phi_H|||_H \lesssim |||\Phi^{\mathrm{ms}}|||_h.$$

Consequently we get

$$|||Q_h(\Phi_H)|||_h \leq |||\Phi^{\mathrm{ms}}|||_h + |||\Phi_H|||_h \lesssim (1 + H^{(1-m)/2})|||\Phi^{\mathrm{ms}}|||_h. \tag{46}$$

Thus

$$\begin{aligned}
a_h\left(\Phi^{\mathrm{ms}},\Phi_H\right) &= a_h(\Phi^{\mathrm{ms}},\Phi^{\mathrm{ms}}) - a_h\left(\Phi^{\mathrm{ms}},Q_h(\Phi_H)\right)\\
&\geq \alpha|||\Phi^{\mathrm{ms}}|||_h^2 - a_h\left(\Phi^{\mathrm{ms}},Q_h(\Phi_H)\right)\\
&= \alpha|||\Phi^{\mathrm{ms}}|||_h^2 - a_h\left(Q_h(\Phi_H)-Q_h^{\Omega}(\Phi_H),Q_h(\Phi_H)\right)\\
&\geq \alpha|||\Phi^{\mathrm{ms}}|||_h^2 - \left|\left|\left|Q_h(\Phi_H)-Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h\ |||Q_h(\Phi_H)|||_h\\
&\overset{(28)}{\geq} \alpha|||\Phi^{\mathrm{ms}}|||_h^2 - \left|\left|\left|Q_h(\Phi_H)-Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h\ (1+H^{(1-m)/2})|||\Phi^{\mathrm{ms}}|||_h.
\end{aligned} \tag{47}$$

Using

$$\begin{aligned}
\left|\left|\left|Q_h(\Phi_H)-Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h &\leq C(1/H)k^{d/2}\theta^k\left|\left|\left|\Phi_H+Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h\\
&\leq C(1/H)k^{d/2}\theta^k\left(|||\Phi^{\mathrm{ms}}|||_h+\left|\left|\left|Q_h(\Phi_H)-Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h\right)\\
&\leq CH^{n-1}\left(|||\Phi^{\mathrm{ms}}|||_h+\left|\left|\left|Q_h(\Phi_H)-Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h\right)
\end{aligned}$$

we obtain that we have for small enough H

$$\left|\left|\left|Q_h(\Phi_H)-Q_h^{\Omega}(\Phi_H)\right|\right|\right|_h \lesssim H^{n-1}|||\Phi^{\mathrm{ms}}|||_h.$$

Inserting this into (47) gives us

$$\begin{aligned}
a_h\left(\Phi^{\mathrm{ms}},\Phi_H\right) &\geq \alpha|||\Phi^{\mathrm{ms}}|||_h^2 - (1+H^{(1-m)/2})H^{n-1}|||\Phi^{\mathrm{ms}}|||_h^2\\
&\geq C_1(\alpha - C_2H)|||\Phi^{\mathrm{ms}}|||_h^2.
\end{aligned}$$

If H is small enough so that $(\alpha - C_2H)$ is positive, the stability estimate $|||\Phi_H|||_H \lesssim |||\Phi^{\mathrm{ms}}|||_h$ concludes the inf-sup estimate. □

Acknowledgments We would like to thank the anonymous referees for their valuable comments and their constructive feedback on the original manuscript which helped us to improve this article.

References

1. Abdulle, A.: On a priori error analysis of fully discrete heterogeneous multiscale FEM. Multiscale Model. Simul. **4**(2), 447–459 (2005)
2. Abdulle, A., E,W., Engquist, B., Vanden-Eijnden, E.: The heterogeneous multiscale method. Acta Numer. **21**, 1–87 (2012)
3. Allaire, G.: Homogenization and two-scale convergence. SIAM J. Math. Anal. **23**(6), 1482–1518 (1992)
4. Aziz, K., Settari, A.: Petroleum Reservoir Simulation. Applied Science Publishers, London (1997)
5. Babuska, I., Lipton, R.: Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. Multiscale Model. Simul. **9**(1), 373–406 (2011)
6. Bank, R.E., Dupont, T.: An optimal order process for solving finite element equations. Math. Comput. **36**(153), 35–51 (1981)
7. Bank, R.E., Yserentant, H.: On the H^1-stability of the L_2-projection onto finite element spaces. Numer. Math. **126**(2), 361–381 (2014)

8. Bush, L., Ginting, V., Presho, M.: Application of a conservative, generalized multiscale finite element method to flow models. J. Comput. Appl. Math. **260**, 395–409 (2014)
9. Carstensen, C.: Quasi-interpolation and a posteriori error analysis in finite element methods. M2AN. Math. Model. Numer. Anal. **33**(6), 1187–1202 (1999)
10. Carstensen, C., Verfürth, R.: Edge residuals dominate a posteriori error estimates for low order finite element methods. SIAM J. Numer. Anal. **36**(5), 1571–1587 (1999)
11. Efendiev, Y., Ginting, V., Hou, T., Ewing, R.: Accurate multiscale finite element methods for two-phase flow simulations. J. Comput. Phys. **220**(1), 155–174 (2006)
12. Efendiev, Y., Hou, T., Ginting, V.: Multiscale finite element methods for nonlinear problems and their applications. Commun. Math. Sci. **2**(4), 553–589 (2004)
13. Elfverson, D., Georgoulis, E.H., Målqvist, A.: An adaptive discontinuous Galerkin multiscale method for elliptic problems. Multiscale Model. Simul. **11**(3), 747–765 (2013)
14. Elfverson, D., Georgoulis, E.H., Målqvist, A., Peterseim, D.: Convergence of a discontinuous Galerkin multiscale method. SIAM J. Numer. Anal. **51**(6), 3351–3372 (2013)
15. Gaspoz, F.D., Heine, C.-J., Siebert, K.G.: Optimal grading of the newest vertex bisection and H^1-stability of the L^2-projection. SimTech Universität, Stuttgart (2014)
16. Ginting, V.: Analysis of two-scale finite volume element method for elliptic problem. J. Numer. Math. **12**(2), 119–141 (2004)
17. Gloria, A.: An analytical framework for the numerical homogenization of monotone elliptic operators and quasiconvex energies. Multiscale Model. Simul. **5**(3), 996–1043 (2006)
18. Gloria, A.: Reduction of the resonance error-part 1: approximation of homogenized coefficients. Math. Models Methods Appl. Sci. **21**(8), 1601–1630 (2011)
19. Henning, P., Målqvist, A.: Localized orthogonal decomposition techniques for boundary value problems. SIAM J. Sci. Comput. **36**(4), A1609–A1634 (2014)
20. Henning, P., Målqvist, A., Peterseim, D.: A localized orthogonal decomposition method for semi-linear elliptic problems. ESAIM Math. Model. Numer. Anal. **48**(5), 1331–1349 (2014)
21. Henning, P., Målqvist, A., Peterseim, D.: Two-level discretization techniques for ground state computations of Bose–Einstein condensates. SIAM J. Numer. Anal. **52**(4), 1525–1550 (2014)
22. Henning, P., Morgenstern, P., Peterseim, D.: Multiscale partition of unity. In: Meshfree Methods for Partial Differential Equations VII. Lecture Notes in Computational Science and Engineering, vol. 100. Springer, Berlin (2015)
23. Henning, P., Ohlberger, M.: The heterogeneous multiscale finite element method for elliptic homogenization problems in perforated domains. Numer. Math. **113**(4), 601–629 (2009)
24. Henning, P., Peterseim, D.: Oversampling for the multiscale finite element method. Multiscale Model. Simul. **11**(4), 1149–1175 (2013)
25. Hou, T.Y., Wu, X.-H.: A multiscale finite element method for elliptic problems in composite materials and porous media. J. Comput. Phys. **134**(1), 169–189 (1997)
26. Hou, T.Y., Wu, X.-H., Zhang, Y.: Removing the cell resonance error in the multiscale finite element method via a Petrov–Galerkin formulation. Commun. Math. Sci. **2**(2), 185–205 (2004)
27. Hughes, T.J.R.: Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. Comput. Methods Appl. Mech. Eng. **127**(1–4), 387–401 (1995)
28. Hughes, T.J.R., Feijóo, G.R., Mazzei, L., Quincy, J.-B.: The variational multiscale method—a paradigm for computational mechanics. Comput. Methods Appl. Mech. Eng. **166**(1–2), 3–24 (1998)
29. Karkulik, M., Pfeiler, C.-M., Praetorius, D.: L^2-orthogonal projections onto finite elements on locally refined meshes are H^1-stable (2013). arXiv:1307.0917
30. Kröner, D.: Numerical schemes for conservation laws. In: Wiley-Teubner Series Advances in Numerical Mathematics. Wiley, Chichester (1997)
31. Larson, M.G., Målqvist, A.: Adaptive variational multiscale ethods based on a posteriori error estimation: duality techniques for elliptic problems. In: Multiscale Methods in Science and Engineering. Lecture Notes in Computational Science and Engineering, vol. 44, pp. 181–193. Springer, Berlin (2005)
32. LeFloch, P.G.: Hyperbolic systems of conservation laws. In: Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2002). (The theory of classical and nonclassical shock waves)
33. Målqvist, A.: Multiscale methods for elliptic problems. Multiscale Model. Simul. **9**(3), 1064–1086 (2011)
34. Målqvist, A., Peterseim, D.: Computation of eigenvalues by numerical upscaling. Numer. Math. (2014). doi:10.1007/s00211-014-0665-6

35. Målqvist, A., Peterseim, D.: Localization of elliptic multiscale problems. Math. Comput. **83**(290), 2583–2603 (2014)
36. Ohlberger, M.: A posteriori error estimates for the heterogeneous multiscale finite element method for elliptic homogenization problems. Multiscale Model. Simul. **4**(1), 88–114 (2005)
37. Owhadi, H., Zhang, L.: Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast. Multiscale Model. Simul. **9**(4), 1373–1398 (2011)
38. Owhadi, H., Zhang, L., Berlyand, L.: Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization. ESAIM Math. Model. Numer. Anal. **48**(2), 517–552 (2014)
39. van der Vorst, H.A.: Computational methods for large eigenvalue problems. In: Handbook of Numerical Analysis, vol. VIII, pp. 3–179. North-Holland, Amsterdam (2002)
40. Weinan, E., Engquist, B.: The heterogeneous multiscale methods. Commun. Math. Sci. **1**(1), 87–132 (2003)