LOCALIZATION OF ELLIPTIC MULTISCALE PROBLEMS

AXEL MÅLQVIST AND DANIEL PETERSEIM

ABSTRACT. This paper constructs a local generalized finite element basis for elliptic problems with heterogeneous and highly varying coefficients. The basis functions are solutions of local problems on vertex patches. The error of the corresponding generalized finite element method decays exponentially with respect to the number of layers of elements in the patches. Hence, on a uniform mesh of size H, patches of diameter $H \log(1/H)$ are sufficient to preserve a linear rate of convergence in H without pre-asymptotic or resonance effects. The analysis does not rely on regularity of the solution or scale separation in the coefficient. This result motivates new and justifies old classes of variational multiscale methods.

1. INTRODUCTION

This paper considers the numerical solution of second order elliptic problems with strongly heterogeneous and highly varying (non-periodic) coefficients. The heterogeneities and oscillations of the coefficient may appear on several non-separated scales. It is well known that classical polynomial based finite element methods perform arbitrarily badly for such problems; see e.g. [4]. To overcome this lack of performance, many methods that are based on general (non-polynomial) ansatz functions have been developed. Early works [1, 2], that essentially apply to one-dimensional problems, have been generalized to the multi-dimensional case in several ways during the last fifteen years; see e.g. [7, 13, 14]. In these methods the problem is split into coarse and (possibly several) fine scales. The fine scale effect on the coarse scale is either computed numerically or modeled analytically. The resulting modified coarse problem can then be solved numerically and its solution contains crucial information from the fine scales. Although many of these approaches show promising results in practice, their convergence analysis usually assumes certain periodicity and scale separation.

For problems with general L^∞ coefficient, the paper [3] gives error bounds for a generalized finite element method that involves the solutions of local eigenvalue problems. The construction in [6, 19] depends only on the solution of the original problem on certain subdomains. However, the size of these subdomains strongly depends on the mesh size. This dependence is suboptimal with respect to the

theoretical statement given in [12], that is, for any shape regular mesh of size H there exist $\mathcal{O}\left((\log(1/H))^{d+1}\right)$ local (non-polynomial) basis functions per nodal point such that the error of the corresponding Galerkin solution u_H satisfies the estimate $\|u - u_H\|_{H^1(\Omega)} \le C_g H$ with a constant C_g that depends on the right-hand side g and the global bounds of the diffusion coefficient but not on its variations. The derivation in [12] is not constructive in the sense that it involves the solution of the (global) original problem with specific right-hand sides.

In this paper, we show that such a (quasi-)optimal basis can indeed be constructed by solving only local problems on element patches. We use a modified nodal basis similar to the one presented in [16] and prove that these basis functions decay exponentially away from the node they are associated with. This exponential decay justifies an approximation using localized patches.

The precise setting of the paper is as follows. Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with polygonal boundary and let the diffusion matrix $A \in L^\infty\left(\Omega, \mathbb{R}^{d \times d}_{\text{sym}}\right)$ be uniformly elliptic:

$$0 < \alpha(A, \Omega) := \operatorname*{ess\,inf}_{x \in \Omega} \inf_{v \in \mathbb{R}^d \setminus \{0\}} \frac{(A(x)v) \cdot v}{v \cdot v},$$

(1.1)

$$\infty > \beta(A, \Omega) := \operatorname*{ess\,sup}_{x \in \Omega} \sup_{v \in \mathbb{R}^d \setminus \{0\}} \frac{(A(x)v) \cdot v}{v \cdot v}.$$

Given $g \in L^2(\Omega)$, we seek $u \in V := H^1_0(\Omega)$ such that

$$(1.2) \qquad a(u, v) := \int_\Omega (A \nabla u) \cdot \nabla v = \int_\Omega gv =: G(v) \quad \text{for all } v \in V.$$

The bilinear form a is symmetric, coercive, bounded, and hence, (1.2) has a unique solution.

The main result of this paper (cf. Theorem 3.6) shows that the error $u - u^{\text{ms}}_{H,k}$ of the generalized finite element method, which is based on our new (local) basis functions mentioned above, is bounded as follows

$$\|A^{1/2} \nabla(u - u^{\text{ms}}_{H,k})\|_{L^2(\Omega)} \le C_g H;$$

H being the mesh size of the underlying coarse finite element mesh and $k \approx \log(1/H)$ referring to the number of layers of coarse elements that form the support of the localized basis functions. This estimate shows that our new numerical upscaling procedure is reliable beyond strong assumptions like periodicity and scale separation. Moreover, the error bound is stable with respect to perturbations arising from the discretization of the local problems. These results give a theoretical foundation for numerous previous experiments where exponential decay of a similar modified basis have been noticed; see e.g. [18].

The outline of the paper is as follows. In Section 2, we derive a set of local basis functions and define the corresponding multiscale finite element method. The error analysis is done in Section 3. Section 4 is devoted to the discretization of the local problems. Section 5 presents numerical experiments, and Section 6 discusses the application of this theory to state-of-the-art multiscale methods.

2. Local basis

In this section, we design a set of local basis functions for the multiscale problem under consideration. The construction is based on a regular (in the sense of [10]) finite element mesh $\mathcal{T}_H$ of Ω into closed triangles ($d = 2$) or tetrahedra ($d = 3$). Subsection 2.1 recalls the classical nodal basis with respect to $\mathcal{T}_H$ and demonstrates its lack of approximation properties. Subsection 2.2 introduces a quasi-interpolation operator used in the construction of the new basis. Subsection 2.3 defines a modified (coefficient dependent) nodal basis and analyzes its approximation properties. This basis is then localized in Subsection 2.4.

2.1. Classical nodal basis.

Let $H : \overline{\Omega} \to \mathbb{R}_{>0}$ denote the $\mathcal{T}_H$-piecewise constant mesh size function with $H|_T = \text{diam}(T) =: H_T$ for all $T \in \mathcal{T}_H$. The mesh size may vary in space. In practical applications, the mesh $\mathcal{T}_H$ (resp., its size H) shall be determined by the accuracy which is desired or the computational capacity that is available but *not* by the scales of the coefficient.

The classical (conforming) P_1 finite element space is given by

$$(2.1) \qquad S_H := \{v \in C^0(\bar{\Omega}) \mid \forall T \in \mathcal{T}_H, v|_T \text{ is a polynomial of total degree} \leq 1\}.$$

Let $V_H := S_H \cap V$ denote the space of finite element functions that match the homogeneous Dirichlet boundary conditions. Let $\mathcal{N}$ denote the set of interior vertices of $\mathcal{T}_H$. For every vertex $x \in \mathcal{N}$, let $\lambda_x \in S_H$ denote the corresponding nodal basis function (tent function), i.e.,

$$\lambda_x(x) = 1 \ \text{ and } \ \lambda_x(y) = 0 \quad \text{for all } y \neq x \in \mathcal{N}.$$

These nodal basis functions form a basis of V_H. The availability of such a local basis is a key property of any finite element method and ensures that the resulting system of linear algebraic equations is sparse.

The (unique) Galerkin approximation $u_H \in V_H$ satisfies

$$(2.2) \qquad\qquad\qquad a(u_H, v) = G(v) \quad \text{ for all } v \in V_H.$$

The above method (2.2) is optimal with respect to the energy norm $|||\cdot||| := |||\cdot|||_\Omega := \|A^{1/2}\nabla \cdot \|_{L^2(\Omega)}$ on V which is induced by a,

$$(2.3) \qquad\qquad\qquad |||u - u_H||| = \min_{v_H \in V_H} |||u - v_H|||.$$

Assuming that the solution u is smooth, the combination of (2.3) and standard interpolation error estimates yields the standard a priori error estimate

$$|||u - u_H||| \leq C\|H\|_{L^\infty(\Omega)}\|\nabla^2 u\|_{L^2(\Omega)}.$$

This estimate states linear convergence of the classical finite element method (2.2) as the maximal mesh width tends to zero. However, the regularity assumption is not realistic for the problem class under consideration. Moreover, even if the coefficient is smooth, it may oscillate rapidly, say at frequency ε^{-1} for some small parameter ε. In this case, the asymptotic result is useless because $\nabla^2 u$ may oscillate at the same scale, a fact that is hidden in the constant $\|\nabla^2 u\|_{L^2(\Omega)} \approx \varepsilon^{-1}$. Unless $H \lesssim \varepsilon$, the above finite element space is unable to capture the behavior of the solution neither on the microscopic nor on the macroscopic level. In what follows, we present a new method that resolves this issue.

2.2. **Quasi-interpolation.** The key tool in our construction will be some bounded linear surjective (quasi-) interpolation operator $\mathfrak{I}_H : V \to V_H$. The choice of this operator is not unique and a different choice might lead to a different multiscale method. We have in mind the following modification of Clément's interpolation [11] which is presented and analyzed in [9, Section 6]. Given $v \in V$, $\mathfrak{I}_H v :=$ $\sum_{x \in \mathcal{N}} (\mathfrak{I}_H v)(x) \lambda_x$ defines a (weighted) Clément interpolant with nodal values

$$(2.4) \qquad (\mathfrak{I}_H v)(x) := \left(\int_\Omega v \lambda_x \right) / \left(\int_\Omega \lambda_x \right)$$

for $x \in \mathcal{N}$. The nodal values are weighted averages of the function over nodal patches $\omega_x := \operatorname{supp} \lambda_x$. Since the summation is taken only with respect to interior vertices $\mathcal{N}$, this operator matches homogeneous Dirichlet boundary conditions.

Recall the (local) approximation and stability properties of the interpolation operator $\mathfrak{I}_H$ [9, Section 6]: There exists a generic constant $C_{\mathfrak{I}_H}$ such that for all $v \in V$ and for all $T \in \mathcal{T}_H$ it holds that

$$(2.5.\text{a}) \qquad H_T^{-1} \|v - \mathfrak{I}_H v\|_{L^2(T)} + \|\nabla(v - \mathfrak{I}_H v)\|_{L^2(T)} \leq C_{\mathfrak{I}_H} \|\nabla v\|_{L^2(\omega_T)},$$

where $\omega_T := \bigcup \{ K \in \mathcal{T}_H \mid T \cap K \neq \emptyset \}$. The constant $C_{\mathfrak{I}_H}$ depends on the shape regularity parameter ρ of the finite element mesh $\mathcal{T}_H$ (see (3.1) below) but not on H_T.

Note that the above interpolation operator is not a projection, i.e., $v_H \in V_H$ does not equal its interpolation $\mathfrak{I}_H v_H$ in general. However, the particular choice gives rise to the following lemma.

Lemma 2.1. *There exists a generic constant $C'_{\mathfrak{I}_H}$ which only depends on ρ but not on the local mesh size H, such that for all $v_H \in V_H$ there exists $v \in V$ with the properties*

$$(2.5.\text{b}) \qquad \mathfrak{I}_H(v) = v_H, \qquad \|\nabla v\| \leq C'_{\mathfrak{I}_H} \|\nabla v_H\|, \quad \text{and} \quad \operatorname{supp} v \subset \operatorname{supp} v_H.$$

Proof. For every nodal basis function λ_x, $x \in \mathcal{N}$, there is some $b_x \in H_0^1(\omega_x)$ such that $\mathfrak{I}_H(b_x) = \lambda_x$ and $\|\nabla b_x\| \leq C''_{\mathfrak{I}_H} \|\nabla \lambda_x\|$ with some constant $C''_{\mathfrak{I}_H}$ that does not depend on x and H. For example, b_x may be chosen as a standard cubic element bubble on an arbitrary element $T \subset \omega_x$ or a quadratic edge/face bubble related to an arbitrary edge/face of $\mathcal{T}_H$ interior to ω_x. One might as well choose b_x to be nodal interpolation of those bubbles in a finite element space that correponds to some uniform refinement of $\mathcal{T}_H$.

Given $v_H = \sum_{x \in \mathcal{N}} v_H(x) \lambda_x \in V_H$, $v := v_H + \sum_{x \in \mathcal{N}} (v_H(x) - (\mathfrak{I}_H v_H)(x)) b_x \in V$ has the desired properties (for suitably chosen b_x). The interpolation and support properties are obvious. The stability follows from

$$
\begin{aligned}
\|\nabla v\|^2 &\leq C \left(\|\nabla v_H\|^2 + \sum_{x \in \mathcal{N}} |v_H(x) - (\mathfrak{I}_H v_H)(x)|^2 \|\nabla b_x\|^2 \right) \\
&\leq C \left(\|\nabla v_H\|^2 + C''^2_{\mathfrak{I}_H} \sum_{x \in \mathcal{N}} |v_H(x) - (\mathfrak{I}_H v_H)(x)|^2 \|\nabla \lambda_x\|^2 \right) \\
&\leq C \left(\|\nabla v_H\|^2 + C' C''^2_{\mathfrak{I}_H} \sum_{T \in \mathcal{T}_H} \|v_H - \mathfrak{I}_H v_H\|^2_{L^2(T)} H_T^{-2} \right) \\
&\leq C \left(\|\nabla v_H\|^2 + C' C^2_{\mathfrak{I}_H} C''^2_{\mathfrak{I}_H} \sum_{T \in \mathcal{T}_H} \|\nabla v_H\|^2_{L^2(\omega_T)} \right) \\
&\leq C'^2_{\mathfrak{I}_H} \|\nabla v_H\|^2,
\end{aligned}
$$

where we use $\|\nabla\lambda_x\|^2 \approx |\operatorname{supp}\lambda_x|^{(d-2)}$, the inverse inequality $\|v_H - \mathfrak{I}_H v_H\|^2_{L^\infty(T)} \lesssim H_T^{-d}\|v_H - \mathfrak{I}_H v_H\|^2_{L^2(T)}$, and (2.5.a). $\qquad\square$

In the forthcoming derivation of our method, the interpolation operator (2.4) may be replaced by any linear bounded surjective operator that satisfies (2.5.a)–(2.5.b). Hereby, (2.5.b) may be relaxed in the sense that $\operatorname{supp} v$ is not necessarily a subset of $\operatorname{supp} v_H$ but that $\operatorname{supp} v \setminus \operatorname{supp} v_H$ covers at most a fixed (small) number of element layers about $\operatorname{supp} v_H$.

2.3. Multiscale splitting and modified nodal basis.

Let $\mathfrak{I}_H : V \to V_H$ be a quasi-interpolation operator according to the previous subsection. Then the kernel of $\mathfrak{I}_H$,

$$V^{\mathrm{f}} := \{v \in V \mid \mathfrak{I}_H v = 0\},$$

represents the microscopic features of V, i.e., all features that are not captured by V_H. Given $v \in V_H$, define $\mathfrak{F}v \in V^{\mathrm{f}}$ by

$$a(\mathfrak{F}v, w) = a(v, w) \quad \text{for all } w \in V^{\mathrm{f}}.$$

The finescale projection operator $\mathfrak{F} : V_H \to V^{\mathrm{f}}$ leads to an orthogonal splitting with respect to the scalar product a:

$$V = V_H^{\mathrm{ms}} \oplus V^{\mathrm{f}} \quad \text{where} \quad V_H^{\mathrm{ms}} := (V_H - \mathfrak{F}V_H).$$

Hence, any function $u \in V$ can be decomposed into $u_H^{\mathrm{ms}} \in V_H^{\mathrm{ms}}$ and $u^{\mathrm{f}} \in V^{\mathrm{f}}$, $u = u_H^{\mathrm{ms}} + u^{\mathrm{f}}$, with $a(u_H^{\mathrm{ms}}, u^{\mathrm{f}}) = 0$. Since $\dim V_H^{\mathrm{ms}} = \dim V_H$, the space V_H^{ms} can be regarded as a modified coarse space. The superscript "ms" abbreviates "multiscale" and indicates that V_H^{ms}, in addition, contains fine scale information. The corresponding Galerkin approximation $u_H^{\mathrm{ms}} \in V_H^{\mathrm{ms}}$ satisfies

$$(2.6) \qquad a(u_H^{\mathrm{ms}}, v) = G(v) \quad \text{for all } v \in V_H^{\mathrm{ms}}.$$

The error $(u - u_H^{\mathrm{ms}})$ of the above method (2.6) is analyzed in Section 3.1.

Finally, we shall introduce a basis of V_H^{ms}. The image of the nodal basis function λ_x under the fine scale projection $\mathfrak{F}$ is denoted by $\phi_x = \mathfrak{F}\lambda_x \in V^{\mathrm{f}}$, i.e., ϕ_x satisfies the corrector problem

$$(2.7) \qquad a(\phi_x, w) = a(\lambda_x, w) \quad \text{for all } w \in V^{\mathrm{f}}.$$

We emphasize that the corrector problem is posed in the fine scale space V^{f}, i.e., test and trial functions satisfy the constraint that their interpolation with respect to the coarse mesh vanishes.

A basis of V_H^{ms} is then given by the modified nodal basis

$$(2.8) \qquad \{\lambda_x - \phi_x \mid x \in \mathcal{N}\}.$$

In general, the corrections ϕ_x of nodal basis functions λ_x, $x \in \mathcal{N}$, have global support, a fact which limits the practical use of the modified basis (2.8) and the corresponding method (2.6).

2.4. **Localization.** In Section 3.2, we will show that the correction ϕ_x decays exponentially fast away from x. Hence, simple truncation of the corrector problems to local patches of coarse elements yields localized basis functions with good approximation properties.

Let $k \in \mathbb{N}$. Define nodal patches of k-th order $\omega_{x,k}$ about $x \in \mathcal{N}$ by

$$
\begin{aligned}
(2.9) \qquad & \omega_{x,1} := \operatorname{supp} \lambda_x = \operatorname{int}\left(\bigcup\{T \in \mathcal{T}_H \mid x \in T\}\right), \\
& \omega_{x,k} := \operatorname{int}\left(\bigcup\{T \in \mathcal{T}_H \mid T \cap \overline{\omega}_{x,k-1} \neq \emptyset\}\right), \quad k = 2, 3, 4 \dots .
\end{aligned}
$$

Define localized finescale spaces $V^{\mathrm{f}}(\omega_{x,k}) := \{v \in V^{\mathrm{f}} \mid v|_{\Omega \setminus \omega_{x,k}} = 0\}$, $x \in \mathcal{N}$, by intersecting V^{f} with those functions that vanish outside the patch $\omega_{x,k}$. The solutions $\phi_{x,k} \in V^{\mathrm{f}}(\omega_{x,k})$ of

$$
(2.10) \qquad a(\phi_{x,k}, w) = a(\lambda_x, w) \quad \text{for all } w \in V^{\mathrm{f}}(\omega_{x,k}),
$$

are approximations of ϕ_x from (2.7) with local support.

We define localized multiscale finite element spaces

$$
(2.11.\mathrm{a}) \qquad V_{H,k}^{\mathrm{ms}} = \operatorname{span}\{\lambda_x - \phi_{x,k} \mid x \in \mathcal{N}\} \subset V.
$$

The corresponding multiscale approximation of (1.2) reads: find $u_{H,k}^{\mathrm{ms}} \in V_{H,k}^{\mathrm{ms}}$ such that

$$
(2.11.\mathrm{b}) \qquad a(u_{H,k}^{\mathrm{ms}}, v) = G(v) \quad \text{for all } v \in V_{H,k}^{\mathrm{ms}}.
$$

Note that $\dim V_{H,k}^{\mathrm{ms}} = |\mathcal{N}| = \dim V_H$, i.e., the number of degrees of freedom of the proposed method (2.11) is the same as for the classical method (2.2). The basis functions of the multiscale method have local support. The overlap is proportional to the parameter k. The error analysis of Section 3.2 suggests to choose $k \approx \log \frac{1}{H}$.

Remark 2.2. The localized modified basis functions could be localized further to vertex patches ω_x, $x \in \mathcal{N}$, by simply multiplying them with the classical nodal basis functions; for any $x \in \mathcal{N}$ and any $y \in \mathcal{N} \cap \omega_{x,k}$, define $\phi_x^y := \lambda_y \phi_{x,k}$. The generalized finite element space which is spanned by those $\mathcal{O}\left((\log(1/H))^d\right)$ local basis functions per vertex has similar approximation properties as $V_{H,k}^{\mathrm{ms}}$ (see [5]).

3. Error analysis

This section analyzes the proposed multiscale method in two steps. First, Subsection 3.1 presents an error bound for the idealized method (2.6). Then, Subsection 3.2 bounds the error of truncation to local patches and proves the main result, that is, an error bound for the multiscale method (2.11).

As usual, the error analysis depends on the constant $\rho > 0$ which represents shape regularity of the finite element mesh $\mathcal{T}_H$;

$$
(3.1) \qquad \rho := \max_{T \in \mathcal{T}_H} \rho_T \quad \text{with} \quad \rho_T := \frac{\operatorname{diam} B_T}{\operatorname{diam} T} \ \text{ for } T \in \mathcal{T}_H,
$$

where B_T denotes the largest ball contained in T.

3.1. **Discretization error.**

Lemma 3.1. *Let $u \in V$ solve (1.2) and $u_H^{\mathrm{ms}} \in V_H^{\mathrm{ms}}$ solve (2.6). Then it holds that*

$$|||u - u_H^{\mathrm{ms}}||| \leq C_{\mathrm{ol}}^{1/2} C_{\mathfrak{I}_H} \alpha^{-1/2} \|Hg\|_{L^2(\Omega)}$$

with constants C_{ol} and $C_{\mathfrak{I}_H}$ that only depend on ρ.

Proof. Recall the (local) approximation and stability properties (2.5.a) of the interpolation operator $\mathfrak{I}_H$. Due to the splitting from Section 2.3, it holds that $u - u_H^{\mathrm{ms}} = u^{\mathrm{f}}$. Since $\mathfrak{I}_H u^{\mathrm{f}} = 0$, the application of (2.5.a) and Young's inequality yield

$$|||u^{\mathrm{f}}|||^2 = G(u^{\mathrm{f}}) \leq \sum_{T \in \mathcal{T}_H} \|g\|_{L^2(T)} \|u^{\mathrm{f}} - \mathfrak{I}_H u^{\mathrm{f}}\|_{L^2(T)}$$

$$\leq \frac{C_{\mathfrak{I}_H}^2}{2\epsilon\alpha} \|Hg\|_{L^2(\Omega)}^2 + \frac{\epsilon}{2} \sum_{T \in \mathcal{T}_H} \|A^{1/2}\nabla u^{\mathrm{f}}\|_{L^2(\omega_T)}^2$$

for any $\epsilon > 0$. Note that there exists a constant $C_{\mathrm{ol}} > 0$ that only depends on ρ such that the number of elements covered by ω_T is uniformly (w.r.t. T) bounded by C_{ol}. The choice $\epsilon = C_{\mathrm{ol}}^{-1}$ concludes the proof. $\square$

Remark 3.2. Substituting $\mathfrak{I}_H$ by the modified Clément interpolation operator presented in [8] allows one to improve the error estimate in Lemma (3.1). The term $\|Hg\|_{L^2(\Omega)}$ can be replaced by data oscillations $(\sum_{x \in \mathcal{N}} \|H(g - g_x)\|_{L^2(\omega_x)}^2)^{1/2}$ with some weighted averages g_x of g on ω_x, $x \in \mathcal{N}$; we refer to [8, Section 2] for details. Additional smoothness of the right-hand side $g \in H^1(\Omega)$ then leads to quadratic convergence of the idealized method without localization.

3.2. **Error of localized multiscale FEM.** First, we estimate the error due to truncation to local patches. We will frequently make use of cut-off functions on element patches.

Definition 3.3. For $x \in \mathcal{N}$ and $m < M \in \mathbb{N}$, let $\eta_x^{m,M} : \Omega \to [0,1]$ be a continuous and weakly differentiable function such that

(3.2.a) $$(\eta_x^{m,M})|_{\omega_{x,m}} = 0,$$

(3.2.b) $$(\eta_x^{m,M})|_{\Omega \setminus \omega_{x,M}} = 1, \text{ and}$$

(3.2.c) $$\forall T \in \mathcal{T}_H, \ \|\nabla \eta_x^{m,M}\|_{L^\infty(T)} \leq C_{\mathrm{co}}(M - m)^{-1} H_T^{-1}$$

with some constant C_{co} that only depends on ρ. For example, one may choose $\eta_x^{m,M} \in S_H$ with nodal values

$$\eta_x^{m,M}(x) = 0 \quad \text{for all } x \in \mathcal{N} \cap \omega_m,$$

(3.3) $$\eta_x^{m,M}(x) = 1 \quad \text{for all } x \in \mathcal{N} \cap (\Omega \setminus \omega_{x,M}), \text{ and}$$

$$\eta_x^{m,M}(x) = j(M - m)^{-1} \quad \text{for all } x \in \mathcal{N} \cap \partial\omega_{x,m+j}, \ j = 0, 1, 2, \ldots, M - m.$$

We prove the essential decay property of the corrector functions by some iterative Caccioppoli-type argument. Recall the notation $|||\cdot|||_\omega := \|A^{1/2}\nabla \cdot\|_{L^2(\omega)}$.

Lemma 3.4. *For all $x \in \mathcal{N}$, $k, \ell \geq 2 \in \mathbb{N}$, the estimate*

$$|||\phi_x - \phi_{x,\ell k}||| \leq C_2 \left(\frac{C_1}{\ell}\right)^{\frac{k-2}{2}} |||\phi_x|||_{\omega_{x,\ell}}$$

holds with constants C_1, C_2 that only depend on ρ and β/α but not on x, k, ℓ, or H.

Proof. Let $x \in \mathcal{N}$ and $\ell, k \geq 2 \in \mathbb{N}$. Observe that

$$(3.4) \qquad |||\phi_x - \phi_{x,\ell k}|||^2 \leq |||\phi_x - v|||^2 = |||\phi_x - v|||_{\omega_{x,\ell k}}^2 + |||\phi_x|||_{\Omega \setminus \omega_{x,\ell k}}^2,$$

holds for all $v \in V^{\mathrm{f}}(\omega_{x,\ell k})$ using Galerkin orthogonality.

Let $\zeta_x := 1 - \eta_x^{\ell(k-1)+1,\ell k-1}$ with a cutoff function $\eta_x^{\ell(k-1)+1,\ell k-1}$ as in Definition 3.3. According to (2.5.b), there exists $b_x \in V$ such that $\mathfrak{I}_H(b_x) = \mathfrak{I}_H(\zeta_x \phi_x)$, $|||b_x||| \leq C'_{\mathfrak{I}_H} |||\mathfrak{I}_H(\zeta_x \phi_x)|||$, and $\operatorname{supp}(b_x) \subset \omega_{x,\ell k}$. Hence, $v := \zeta_x \phi_x - b_x \in V^{\mathrm{f}}(\omega_{x,\ell k})$ and

$$|||\phi_x - v|||_{\omega_{x,\ell k}} \leq |||\phi_x - \zeta_x \phi_x|||_{\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)+1}} + |||b_x|||_{\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)}}$$
$$\leq C'_{\mathfrak{I}_H} C_{\mathfrak{I}_H} \left(|||\phi_x|||_{\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)+1}} + \sqrt{\beta} \|\nabla(\zeta_x \phi_x)\|_{L^2(\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)})} \right).$$

Since $\mathfrak{I}_H \phi_x = 0$, the upper bound of the interpolation error (2.5.a) and (3.2.c) yield

$$\|\nabla(\zeta_x \phi_x)\|_{L^2(\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)})}^2$$
$$\leq C_2''' \sum_{T \in \mathcal{T}_H: \, T \subset \overline{\omega}_{x,\ell k} \setminus \omega_{x,\ell(k-1)+1}} \left(H_T^2 \|\nabla \zeta_k\|_{L^\infty(T)}^2 + \|\zeta_k\|_{L^\infty(T)}^2 \right) \|\nabla \phi_x\|_{L^2(T)}^2$$
$$\leq C_2'' \alpha^{-1} |||\phi_x|||_{\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)+1}}^2$$

with $C_2'' := 1 + C_{\mathrm{ol}} C_{\mathrm{co}}^2 C_{\mathfrak{I}_H}^2$. This leads to

$$(3.5) \qquad |||\phi_x - v|||_{\omega_{x,\ell k}} \leq C_2' |||\phi_x|||_{\omega_{x,k\ell} \setminus \omega_{x,(k-1)\ell}},$$

where C_2' depends only on ρ and $\sqrt{\beta/\alpha}$. The combination of (3.4), with $v = \zeta_x \phi_x - b_x$, and (3.5) yields

$$(3.6) \qquad |||\phi_x - \phi_{x,\ell k}||| \leq C_2 |||\phi_x|||_{\Omega \setminus \omega_{x,\ell(k-1)}}.$$

Further estimation of the right-hand side in (3.6) is possible using cut-off functions $\eta_j := \eta_x^{\ell(j-1)+1,\ell j}$ (cf. Definition 3.3), $j = 2, 3, \ldots, k-1$. Observe that

$$(3.7) \qquad \|A^{1/2} \nabla \phi_x\|_{L^2(\Omega \setminus \omega_{x,\ell(k-1)})}^2 \leq \|A^{1/2} \eta_{k-1} \nabla \phi_x\|_{L^2(\Omega)}^2$$
$$= \int_\Omega (A \nabla \phi_x) \cdot \nabla(\eta_{k-1}^2 \phi_x) \, - \, 2 \int_\Omega \eta_{k-1} \phi_x (A \nabla \phi_x) \cdot \nabla \eta_{k-1}.$$

Let, according to Lemma 2.1, $b_{x,(k-1)}$ be chosen such that $\mathfrak{I}_H b_{x,(k-1)} = \mathfrak{I}_H(\eta_{k-1}^2 \phi_x)$. Then $\eta_{k-1}^2 \phi_x - b_{x,(k-1)} \in V^{\mathrm{f}}$. Since $|\operatorname{supp}(\nabla \lambda_x) \cap \operatorname{supp}(\eta_{k-1})| = 0$ and $\operatorname{supp}(\nabla \eta_{k-1}) = \omega_{x,(k-1)\ell} \setminus \omega_{x,(k-2)\ell+1}$, the first term on the right-hand side of (3.7) can be rewritten as

$$(3.8)$$
$$\int_\Omega (A \nabla \phi_x) \cdot \nabla(\eta_{k-1}^2 \phi_x)$$
$$= \int_\Omega (A \nabla \phi_x) \cdot \nabla(\eta_{k-1}^2 \phi_x - b_{x,(k-1)}) + \int_\Omega (A \nabla \phi_x) \cdot \nabla b_{x,(k-1)}$$
$$= \int_\Omega (A \nabla \phi_x) \cdot \nabla b_{x,(k-1)}$$
$$\leq C'_{\mathfrak{I}_H} \sqrt{\beta} |||\phi_x|||_{\omega_{x,(k-1)\ell} \setminus \omega_{x,(k-2)\ell+1}} \|\nabla \mathfrak{I}_H(\eta_{k-1}^2 \phi_x)\|_{L^2(\omega_{x,(k-1)\ell} \setminus \omega_{x,(k-2)\ell+1})}.$$

With $\overline{\eta_T^2} := |T|^{-1} \int_T \eta_{k-1}^2$ we have

$$\|\nabla \mathfrak{I}_H(\eta_{k-1}^2 \phi_x)\|_{L^2(T)} = \|\nabla \mathfrak{I}_H((\eta_{k-1}^2 - \overline{\eta_T^2})\phi_x)\|_{L^2(T)}$$

$$\leq C_{\mathfrak{I}_H} \|\nabla((\eta_{k-1}^2 - \overline{\eta_T^2})\phi_x)\|_{L^2(T)}$$

$$\leq C_{\mathfrak{I}_H} \left(\|\eta_{k-1}^2 - \overline{\eta_T^2}\|_{L^\infty(T)} \|\nabla \phi_x\|_{L^2(T)} + \|\nabla(\eta_{k-1}^2)\|_{L^\infty(T)} \|\phi_x\|_{L^2(T)} \right)$$

$$\leq 2C_{\mathfrak{I}_H} \|\nabla(\eta_{k-1})\|_{L^\infty(T)} \left(\alpha^{-1/2} \operatorname{diam}(T) \, \||\phi_x\||_T + \|\phi_x - \mathfrak{I}_H(\phi_x)\|_{L^2(T)} \right).$$

Thus, the property (3.2.c) of the cutoff function and the upper bound of the interpolation error (2.5.a) yield

$$(3.9) \qquad \||\mathfrak{I}_H(\eta_{k-1}^2 \phi_x)\||_{\omega_{x,(k-1)\ell} \setminus \omega_{x,(k-2)\ell+1}} \leq C_1' \ell^{-1} \|A^{1/2} \nabla \phi_x\|_{L^2(\Omega \setminus \omega_{x,(k-2)\ell})},$$

where C_1' only depends on $C_{\mathfrak{I}_H}$, C_{co}, C_{ol}, and $\sqrt{\beta/\alpha}$. The same arguments allow one to bound the second term on the right-hand side in (3.7),
(3.10)

$$2 \int_\Omega \eta_{k-1} \phi_x (A \nabla \phi_x) \cdot \nabla \eta_{k-1}$$

$$\leq 2 \sum_{T \in \mathcal{T}_H : \, T \subset \overline{\omega}_{x,(k-1)\ell} \setminus \omega_{x,(k-2)\ell+1}} \|\nabla \eta_{k-1}\|_{L^\infty(T)} \|A^{1/2} \nabla \phi_x\|_{L^2(T)} \|A^{1/2} \phi_x\|_{L^2(T)}$$

$$\leq C_1'' \ell^{-1} \|A^{1/2} \nabla \phi_x\|_{L^2(\Omega \setminus \omega_{x,(k-2)\ell})}^2,$$

where C_1'' only depends on $C_{\mathfrak{I}_H}$, C_{co}, and $\sqrt{\beta/\alpha}$. The combination of (3.7)–(3.10) yields

$$(3.11) \qquad \||\phi_x\||_{\Omega \setminus \omega_{x,(k-1)\ell}}^2 \leq C_1 \ell^{-1} \, \||\phi_x\||_{\Omega \setminus \omega_{x,(k-2)\ell}}^2,$$

where $C_1 := C_1' + C_1''$. For $j = k-2, \ldots, 2$, a similar argument (with η_{k-1} replaced by η_j) yields

$$(3.12) \qquad \||\phi_x\||_{\Omega \setminus \omega_{x,j\ell}}^2 \leq C_1 \ell^{-1} \, \||\phi_x\||_{\Omega \setminus \omega_{x,(j-1)\ell}}^2.$$

Starting from (3.11), the successive application of (3.12) for $j = k-2, k-3, \ldots, 2$ proves

$$(3.13) \qquad \||\phi_x\||_{\Omega \setminus \omega_{x,(k-1)\ell}}^2 \leq (C_1 \ell^{-1})^{k-2} \, \||\phi_x\||_{\omega_{x,\ell}}^2.$$

Combining (3.6) and (3.13), we finally obtain the assertion. $\qquad \square$

Lemma 3.5. *There is a constant C_3 that depends only on ρ and β/α, but not on $|\mathcal{N}|$, k, or ℓ such that*

$$\left\| \left| \left\| \sum_{x \in \mathcal{N}} v(x)(\phi_x - \phi_{x,\ell k}) \right\| \right| \right\|^2 \leq C_3 (\ell k)^d \sum_{x \in \mathcal{N}} v^2(x) \, \||\phi_x - \phi_{x,\ell k}\||^2.$$

Proof. For $x \in \mathcal{N}$, let $\zeta_x = 1 - \eta_x^{\ell k+1, \ell k+2}$ (cf. Definition 3.3). By Lemma 2.1 there exists a function $b_x \in V$ such that for any $w \in V^{\mathrm{f}}$ it holds that

$$\mathfrak{I}_H b_x = \mathfrak{I}_H((1 - \zeta_x)w), \operatorname{supp}(b_x) \subset \operatorname{supp}(\mathfrak{I}_H((1-\zeta_x)w)) \subset \omega_{x,\ell k+3} \setminus \omega_{x,\ell k},$$

and

$$\||b_x\||_{\omega_{x,\ell k+3} \setminus \omega_{x,\ell k}} \leq C_{\mathfrak{I}_H}' \, \||\mathfrak{I}_H((1-\zeta_x)w)\||_{\omega_{x,\ell k+3} \setminus \omega_{x,\ell k}}.$$

We note that $w - \zeta_x w - b_x \in V^{\mathrm{f}}$ with support outside $\omega_{x,\ell k}$, i.e., $a(\phi_x, w - \zeta_x w - b_x) = a(\lambda_x, w - \zeta_x w - b_x) = 0$ and $a(\phi_{x,\ell k}, w - \zeta_x w - b_x) = 0$. With $w = \sum_{x \in \mathcal{N}} v(x)(\phi_x - \phi_{x,\ell k}) \in V^{\mathrm{f}}$ we have

$$
\begin{aligned}
|||w|||^2 &= \sum_{x \in \mathcal{N}} v(x)\, a(\phi_x - \phi_{x,\ell k}, \zeta_x w + b_x) \\
&\leq \sqrt{\beta} \sum_{x \in \mathcal{N}} |v(x)|\, |||\phi_x - \phi_{x,\ell k}||| \cdot \|\nabla(\zeta_x w)\|_{L^2(\Omega)} \\
&\quad + \sqrt{\beta} \sum_{x \in \mathcal{N}} |v(x)|\, |||\phi_x - \phi_{x,\ell k}||| \cdot C'_{\mathfrak{I}_H} \|\nabla(\mathfrak{I}_H((1-\zeta_x)w))\|_{L^2(\omega_{x,\ell k+3})} \\
&\leq 2\sqrt{\beta} C'_{\mathfrak{I}_H} C_{\mathfrak{I}_H} \sum_{x \in \mathcal{N}} |v(x)|\, |||\phi_x - \phi_{x,\ell k}||| \cdot \|\nabla(\zeta_x w)\|_{L^2(\Omega)} \\
&\quad + 2\sqrt{\beta} C'_{\mathfrak{I}_H} C_{\mathfrak{I}_H} \sum_{x \in \mathcal{N}} |v(x)|\, |||\phi_x - \phi_{x,\ell k}||| \cdot \|\nabla w\|_{L^2(\omega_{x,\ell k+4})} \\
&\leq 2\sqrt{\beta} C'_{\mathfrak{I}_H} C_{\mathfrak{I}_H} \sum_{x \in \mathcal{N}} |v(x)|\, |||\phi_x - \phi_{x,\ell k}||| \cdot \|(\nabla\zeta_x)(1 - \mathfrak{I}_H)w)\|_{L^2(\omega_{x,\ell k+2})} \\
&\quad + 2\sqrt{\frac{\beta}{\alpha}} C'_{\mathfrak{I}_H} C_{\mathfrak{I}_H} \sum_{x \in \mathcal{N}} |v(x)|\, |||\phi_x - \phi_{x,\ell k}||| \cdot |||w|||_{\omega_{x,\ell k+4}} \\
&\leq 4\sqrt{\frac{\beta}{\alpha}} C'_{\mathfrak{I}_H} C^2_{\mathfrak{I}_H} C_{\mathrm{co}} \sum_{x \in \mathcal{N}} |v(x)|\, |||\phi_x - \phi_{x,\ell k}||| \cdot |||w|||_{\omega_{x,\ell k+4}} \\
&\leq 4\sqrt{\frac{\beta}{\alpha}} C'_{\mathfrak{I}_H} C^2_{\mathfrak{I}_H} C_{\mathrm{co}} C_{\mathrm{ov}}(\ell k)^{d/2} \left(\sum_{x \in \mathcal{N}} v^2(x)\, |||\phi_x - \phi_{x,\ell k}|||^2 \right)^{1/2} |||w|||,
\end{aligned}
$$

where $C_{\mathrm{ov}}(\ell k)^d$ represents an upper bound on the number of patches $\omega_{x,\ell k}$ that overlap a single element in the mesh. The result follows by dividing by $|||w|||$ on both sides. $\qquad\square$

Theorem 3.6. *Let $u \in V$ solve (1.2) and, given $\ell, k \geq 2 \in \mathbb{N}$, let $u^{\mathrm{ms}}_{H,\ell k} \in V^{ms}_{H,\ell k}$ solve (2.11). Then*

$$
\begin{aligned}
|||u - u^{\mathrm{ms}}_{H,\ell k}||| &\leq C_4 \|H_T^{-1}\|_{L^\infty(\Omega)} (\ell k)^{d/2} (C_1/\ell)^{\frac{k-2}{2}} \|g\|_{H^{-1}(\Omega)} \\
&\quad + C^{1/2}_{\mathrm{ol}} C_{\mathfrak{I}_H} \alpha^{-1/2} \|Hg\|_{L^2(\Omega)}
\end{aligned}
$$

holds with C_1 from Lemma 3.4 and a constant C_4 that depends on α, β and ρ but not on H, k, ℓ, g, or u.

Proof. Let $\tilde{u}^{\mathrm{ms}}_{H,\ell k} := \sum_{x \in \mathcal{N}} u^{\mathrm{ms}}_H(x) (\lambda_x - \phi_{x,\ell k})$, where $u^{\mathrm{ms}}_H(x)$, $x \in \mathcal{N}$, are the coefficients in the basis representation of u^{ms}_H. Due to Galerkin orthogonality, Lemma 3.1, Lemma 3.5, and the triangle inequality,

$$
\begin{aligned}
(3.14) \qquad |||u - u^{\mathrm{ms}}_{H,\ell k}||| &\leq |||u - \tilde{u}^{\mathrm{ms}}_{H,\ell k}||| = |||u - u^{\mathrm{ms}}_H + u^{\mathrm{ms}}_H - \tilde{u}^{\mathrm{ms}}_{H,\ell k}||| \\
&\leq C^{1/2}_{\mathrm{ol}} C_{\mathfrak{I}_H} \alpha^{-1/2} \|Hg\|_{L^2(\Omega)} + |||u^{\mathrm{ms}}_H - \tilde{u}^{\mathrm{ms}}_{H,\ell k}|||.
\end{aligned}
$$

The application of Lemma 3.4 yields

$$\left|\left|\left|u_H^{\mathrm{ms}} - \tilde{u}_{H,\ell k}^{\mathrm{ms}}\right|\right|\right|^2 \leq C_3(\ell k)^d \sum_{x\in\mathcal{N}} u_H^{\mathrm{ms}}(x)^2 \left|\left|\left|\phi_x - \phi_{x,\ell k}\right|\right|\right|^2$$

$$\leq C_3(\ell k)^d C_2^2 (C_1/\ell)^{k-2} \sum_{x\in\mathcal{N}} u_H^{\mathrm{ms}}(x)^2 \left|\left|\left|\phi_x\right|\right|\right|_{\omega_{x,\ell}}^2.$$

Furthermore, we have

$$\sum_{x\in\mathcal{N}} u_H^{\mathrm{ms}}(x)^2 \left|\left|\left|\phi_x\right|\right|\right|_{\omega_{x,l}}^2 \leq \beta C_{\mathrm{inv}} \sum_{T\in\mathcal{T}} H_T^{-2} \sum_{x\in T\cap\mathcal{N}} u_H^{\mathrm{ms}}(x)^2 \|\lambda_x\|_{L^2(T)}^2$$

$$\leq \beta C_{\mathrm{inv}}' \sum_{T\in\mathcal{T}} H_T^{-2} \left\| \sum_{x\in T\cap\mathcal{N}} u_H^{\mathrm{ms}}(x)\lambda_x \right\|_{L^2(T)}^2$$

$$= \beta C_{\mathrm{inv}}' \left\| H^{-2} \sum_{x\in\mathcal{N}} u_H^{\mathrm{ms}}(x)\lambda_x \right\|_{L^2(\Omega)}^2$$

$$\leq \beta C_{\mathrm{inv}}' \left(\|H^{-2} u_H^{\mathrm{ms}}\|_{L^2(\Omega)}^2 + \left\| H^{-2} \sum_{x\in\mathcal{N}} u_H^{\mathrm{ms}}(x)(\phi_x - \mathfrak{I}_H\phi_x) \right\|_{L^2(\Omega)}^2 \right)$$

$$\leq \frac{\beta}{\alpha} C_{\mathrm{inv}}' (C_{\mathrm{F}} \|H_T^{-2}\|_{L^\infty(\Omega)} + C_{\mathfrak{I}_H}) \left|\left|\left|u_H^{\mathrm{ms}}\right|\right|\right|^2,$$

where C_{inv} and C_{inv}' depend on ρ and $C_{\mathrm{F}} = C_{\mathrm{F}}(\Omega)$ is the constant from Friedrichs' inequality. This yields

(3.15)
$$\left|\left|\left|u_H^{\mathrm{ms}} - \tilde{u}_{H,\ell k}^{\mathrm{ms}}\right|\right|\right| \leq C_4' \|H_T^{-1}\|_{L^\infty(\Omega)} (\ell k)^{d/2} (C_1/\ell)^{(k-2)/2} \left|\left|\left|u_H^{\mathrm{ms}}\right|\right|\right|$$

$$\leq C_4 \|H_T^{-1}\|_{L^\infty(\Omega)} (\ell k)^{d/2} (C_1/\ell)^{(k-2)/2} \|g\|_{H^{-1}(\Omega)},$$

where C_4 only depends on C_2, C_3, C_{inv}', C_{F}, $C_{\mathfrak{I}_H}$, and $\sqrt{\beta}/\alpha$. The assertion follows readily by combining (3.14) and (3.15). $\qquad\square$

Remark 3.7. The error estimate in Theorem 3.6 contains a factor $\|H^{-1}\|_{L^\infty(\Omega)}$. However, its influence on the total error can be controlled by choosing the localization parameter k proportional to $\log(1/\|H^{-1}\|_{L^\infty(\Omega)})$. For non-uniform meshes, it is recommended to vary the choice of the localization parameter in space according to $k \approx \log \frac{1}{H}$. We neglect this opportunity to avoid overloading the paper.

4. Discretization of the fine scale computations

In this section, we focus on how to compute numerical approximations to the local basis functions $\lambda_x - \phi_{x,\ell k}$ and thereby to the multiscale solution $u_{H,\ell k}^{\mathrm{ms}}$. In order to do this, we need to extend the error analysis of Section 3 to a fully discrete setting. There is a lot of freedom in choosing different finite elements and different refinement strategies; see e.g. [16, 17]. We will focus on a very simple and natural approach. We assume that the local basis functions are computed using subgrids of a fine scale reference mesh, which is a (possibly space adaptive) refinement of the coarse grid $\mathcal{T}_H$.

More precisely, let $\mathcal{T}_h$ be the result of one uniform refinement and several conforming but possibly non-uniform refinements of the coarse mesh $\mathcal{T}_H$. We introduce $h : \overline{\Omega} \to \mathbb{R}_{>0}$ as the $\mathcal{T}_h$-piecewise constant mesh width function with $h_t := h|_t = \mathrm{diam}(t)$ for all $t \in \mathcal{T}_h$. We construct the finite element space

$$S_h := \{v \in C^0(\Omega) \mid \forall t \in \mathcal{T}_h(\Omega), v|_t \text{ is a polynomial of total degree} \leq 1\}.$$

We let $u_h \in V_h := S_h \cap H_0^1(\Omega)$ be the reference solution that satisfies

$$(4.1) \qquad\qquad a(u_h, v) = G(v) \quad \text{for all } v \in V_h.$$

Locally on each patch we let

$$(4.2) \qquad V_h^{\mathrm{f}}(\omega_{x,k}) := V^{\mathrm{f}}(\omega_{x,k}) \cap V_h = \{v \in V_h \mid \mathfrak{I}_H v = 0 \text{ and } v|_{\Omega \setminus \omega_{x,k}} = 0\}.$$

The numerical approximation $\phi_{x,k}^h \in V_h^{\mathrm{f}}(\omega_{x,k})$ of the corrector $\phi_{x,k}^h$ is determined by

$$a(\phi_{x,k}^h, w) = a(\lambda_x, w) \quad \text{for all } w \in V_h^{\mathrm{f}}(\omega_{x,k}).$$

We denote the discrete multiscale finite element space

$$V_{H,k}^{\mathrm{ms},h} = \mathrm{span}\{\lambda_x - \phi_{x,k}^h \mid x \in \mathcal{N}\}.$$

The corresponding discrete multiscale approximation $u_{H,k}^{\mathrm{ms},h} \in V_{H,k}^{\mathrm{ms},h}$ fulfills

$$(4.3) \qquad\qquad a(u_{H,k}^{\mathrm{ms},h}, v) = G(v) \quad \text{for all } v \in V_{H,k}^{\mathrm{ms},h}.$$

Theorem 4.1. *Let $u \in V$ solve (1.2) and let $u_{H,\ell k}^{\mathrm{ms},h} \in V_{H,k}^{\mathrm{ms},h}$ solve (4.3). Then*

$$\left|\left|\left| u - u_{H,\ell k}^{\mathrm{ms},h} \right|\right|\right| \leq \tilde{C}_4 \|H_T^{-1}\|_{L^\infty(\Omega)} (\ell k)^{d/2} (\tilde{C}_1/\ell)^{\frac{k-2}{2}} \|g\|_{H^{-1}(\Omega)}$$

$$+ C_{\mathrm{ol}}^{1/2} C_{\mathfrak{I}_H} \alpha^{-1/2} \|Hg\|_{L^2(\Omega)} + |||u - u_h|||,$$

where $\tilde{C}_4$ only depends on ρ, α and β.

Remark 4.2 (Multiscale splitting by nodal interpolation). Having discretized the fine scale computation, i.e., having replaced the infinite dimensional space V by some finite element space $V_h \subset C^0(\overline{\Omega})$ we are allowed to replace the Clément-type interpolation by classical nodal interpolation. This leads to the variational multiscale method in [18], which is a modification of the method first presented in [17]. Because nodal interpolation satisfies the conditions (2.5.a)–(2.5.b), Theorem 4.1 establishes an a priori error bound for the multiscale method [18]. However, the constant $C_{\mathfrak{I}_H}$ in (2.5.a) depends on the ratio H/h of the discretization scales if $d > 1$ ($C_{\mathfrak{I}_H} \approx \log(H/h)$ in $2d$ and $C_{\mathfrak{I}_H} \approx (H/h)^{-1}$ in $3d$, cf. [21]). Hence, for nodal interpolation, the constants $\tilde{C}_1, \tilde{C}_4$ in Theorem 4.1 depend on H/h in a similar fashion. In 2d this can still be acceptable because the dependence on H/h is only logarithmic.

Remark 4.3 (Estimates for the fine scale error). The finite element space V_h may be replaced by any finite element space that contains V_h, e.g., by piecewise polynomials of higher order. The third part in the error bound in Theorem 4.1 can be bounded in terms of data, mesh parameter h, and polynomial degree using standard a priori error estimates. For example, if $A \in W^{1,\infty}(\Omega)$ (bounded with bounded weak derivative) and ε is the smallest present scale, i.e., $\|\nabla A\|_{L^\infty(\Omega)} \lesssim \varepsilon^{-1}$, the third term in the error bound in Theorem 4.1 may be replaced by the worst case bound $Ch\varepsilon^{-1}$ for a first-order ansatz space V_h (see [20]). It is shown in [20] that for highly varying but smooth coefficient A, higher order ansatz spaces are superior.

Remark 4.4 (Periodic coefficient). Let Ω be some square or cube, $g \in L^2(\Omega)$, let A be smooth and periodic, $A(x) = A(x/\varepsilon)$, with some small scale parameter $\varepsilon > 0$, and let u_ε denote the corresponding solution of (1.2). Choose uniform meshes $\mathcal{T}_H$

and $\mathcal{T}_h$ with $H > \varepsilon > h$ and $k \approx \log(H^{-1})$. With regard to the previous comment, Theorem 4.1 yields the error bound

$$\left|\left|\left|u_\varepsilon - u_{H,\ell k}^{\mathrm{ms},h}\right|\right|\right| \leq C_g(H + \tfrac{h}{\varepsilon}).$$

With $h \sim \varepsilon H$ the error in the approximation becomes independent of the fine scale oscillations without any so-called resonance effects as they are observed, e.g., in [13]. We emphasize that periodicity can be exploited to reduce the number corrector problems to be solved significantly.

Remark 4.5 (Solution of the local problems). The local problems need to be solved in the spaces $V_h^{\mathrm{f}}(\omega_{x,k})$. This is a standard finite element space with the additional constraint that the trial and test functions should have no component in V_H. In practice this constraint is realized using Lagrange multipliers.

The resulting coarse scale system of equations is of the same size as the original problem, $\dim(V_{H,k}^{\mathrm{ms},h}) = \dim V_H$ and it is still sparse. The number of non-zero entries will be larger and depend on k. Note, however, that the non-zero entries in the stiffness matrix decay exponentially away from the diagonal.

Proof of Theorem 4.1. We use the triangle inequality

$$\left|\left|\left|u - u_{H,\ell k}^{\mathrm{ms},h}\right|\right|\right| \leq |||u - u_h||| + \left|\left|\left|u_h - u_H^{\mathrm{ms},h}\right|\right|\right| + \left|\left|\left|u_H^{\mathrm{ms},h} - u_{H,\ell k}^{\mathrm{ms},h}\right|\right|\right|$$

and follow the arguments from the proof of Theorem 3.6 simply replacing V by V_h and using Lemmas 4.6, 4.8, and 4.9 below (discrete versions of Lemmas 3.1, 3.4, and 3.5) to bound the last two terms. $\qquad\square$

Lemma 4.6 (Discrete version of Lemma 3.1). *Let $u_h \in V_h$ solve (4.1) and $u_H^{\mathrm{ms},h} \in V_H^{\mathrm{ms},h}$ solve (4.3) with k large enough so that $\omega_{x,k} = \Omega$ for all $x \in \mathcal{N}$. Then*

$$\left|\left|\left|u_h - u_H^{\mathrm{ms},h}\right|\right|\right| \leq C_{\mathrm{ol}}^{1/2} C_{\mathfrak{J}_H} \alpha^{-1/2} \|Hg\|_{L^2(\Omega)}$$

holds with constants C_{ol} and $C_{\mathfrak{J}_H}$ that only depend on ρ.

Proof. Note that $u_h^{\mathrm{f}} := u_h - u_H^{\mathrm{ms},h}$ is the unique element of $V_h^{\mathrm{f}} := V^{\mathrm{f}} \cap V_h$ such that $a(u_h^{\mathrm{f}}, v) = G(v)$ for all $v \in V_h^{\mathrm{f}}$. The lemma follows from the same arguments in the proof of Lemma 3.1. $\qquad\square$

In the remaining part of this Section, $A \lesssim B$ abbreviates an inequality $A \leq C B$ with some generic constant $0 \leq C < \infty$ that does not depend on the mesh sizes H, h and the localization parameters. The constant may depend on the contrast β/α but not on the geometrical or topological structure of the coefficient A.

To establish discrete versions of Lemmas 3.4 and 3.5 we are facing the technical difficulty that the product of $v \in V_h$ and some cut-off function η from Definition 3.3 is not necessarily an element of V_h. However, the subsequent lemma shows that the product ηv can be approximated sufficiently well by elements from V_h.

Lemma 4.7. *For all $x \in \mathcal{N}$, $M > m \in \mathbb{N}$, and corresponding cut-off function $\eta_x^{m,M}$ defined in (3.3) there exists some $v \in V_h^{\mathrm{f}}(\omega_{x,M+1})$ such that*

$$\left|\left|\left|\eta_x^{m,M}\phi_x^h - v\right|\right|\right| \lesssim \frac{1}{M-m}\left|\left|\left|\phi_x^h\right|\right|\right|_{\omega_{x,M+1}\backslash\omega_{x,m-1}}.$$

Furthermore, the statement also holds if $\eta_x^{m,M}$ is replaced by $1 - \eta_x^{m,M}$ and $v \in V_h^{\mathrm{f}}(\Omega \setminus \omega_{x,m-1})$.

Proof. Let $x \in \mathcal{N}$, $M > m \in \mathbb{N}$ be fixed and define $\eta := \eta_x^{m,M}$. Let $\mathfrak{I}_h : V \cap C(\bar{\Omega}) \to V^h$ be the nodal interpolant with respect to the mesh $\mathcal{T}_h$. Recall its (local) approximation and stability properties

$$\|\nabla(v - \mathfrak{I}_h v)\|_{L^2(t)} \lesssim h_t \|\nabla^2 v\|_{L^2(t)} \quad \text{and} \quad \|\mathfrak{I}_h v\|_{L^2(t)} \lesssim \|v\|_{L^2(t)}$$

for all polynomials v. According to Lemma 2.1, there exists some $b_x \in V^h$ such that $\mathfrak{I}_H(b_x) = \mathfrak{I}_H(\mathfrak{I}_h(\eta \phi_x^h))$, $|||b_x||| \lesssim |||\mathfrak{I}_H(\mathfrak{I}_h(\eta \phi_x^h))|||$, and $\mathrm{supp}(b_x) \subset \omega_{x,M+1} \setminus \omega_{x,m-1}$. Hence, $v := \mathfrak{I}_h(\eta \phi_x^h) - b_x \in V_h^{\mathrm{f}}(\omega_{x,M+1})$. Since $\mathfrak{I}_H \mathfrak{I}_h \bar{\eta}_T \phi_x^h = \bar{\eta}_T \mathfrak{I}_H \phi_x^h = 0$ for $\bar{\eta}_T = |T|^{-1} \int_T \eta$, we get

$$
\begin{aligned}
|||\eta \phi_x^h - v|||^2 &= |||\eta \phi_x^h - \mathfrak{I}_h(\eta \phi_x^h) + b_x|||^2 \\
&\lesssim \sum_{t \in \mathcal{T}_h : t \subset \bar{\omega}_{x,M} \setminus \omega_{x,m}} \|\nabla(\eta \phi_x^h - \mathfrak{I}_h(\eta \phi_x^h))\|_{L^2(t)}^2 + |||\mathfrak{I}_H(\mathfrak{I}_h((\eta - \bar{\eta}_T)\phi_x^h))|||^2 \\
&\lesssim \sum_{t \in \mathcal{T}_h : t \subset \bar{\omega}_{x,M} \setminus \omega_{x,m}} h_t^2 \|\nabla^2(\eta \phi_x^h)\|_{L^2(t)}^2 + \sum_{T \in \mathcal{T}_H : T \subset \bar{\omega}_{x,M+1} \setminus \omega_{x,m-1}} H_T^{-2} \|\mathfrak{I}_h((\eta - \bar{\eta}_T)\phi_x^h)\|_{L^2(T)}^2 \\
&\lesssim \sum_{t \in \mathcal{T}_h : t \subset \bar{\omega}_{x,M} \setminus \omega_{x,m}} h_t^2 \left(\|\nabla^2 \eta\|_{L^\infty(t)}^2 \|\phi_x^h\|_{L^2(t)}^2 + \|\nabla \eta\|_{L^\infty(t)}^2 \|\nabla \phi_x^h\|_{L^2(t)}^2 \right) \\
&\quad + \sum_{T \in \mathcal{T}_H : T \subset \bar{\omega}_{x,M+1} \setminus \omega_{x,m-1}} H_T^{-2} \|\eta - \bar{\eta}_T\|_{L^\infty(T)}^2 \|\phi_x^h\|_{L^2(T)}^2 \\
&\lesssim (M - m)^{-1} |||\phi_x^h|||_{\omega_{x,M+1} \setminus \omega_{x,m-1}}^2
\end{aligned}
$$

using the property (3.2.c) of η and Poincaré's inequality. This proves the first part of the lemma.

The second part concerning $1 - \eta$ follows using the same argument but with $v \in V_h^{\mathrm{f}}(\Omega \setminus \omega_{x,m-1})$. $\qquad \square$

Lemma 4.8 (Discrete version of Lemma 3.4). *For all $x \in \mathcal{N}$, $k, \ell \geq 2 \in \mathbb{N}$ the estimate*

$$|||\phi_x^h - \phi_{x,\ell k}^h||| \leq \tilde{C}_2 \left(\frac{\tilde{C}_1}{\ell} \right)^{\frac{k-2}{2}} |||\phi_x^h|||_{\omega_{x,\ell}}$$

holds with constants $\tilde{C}_1, \tilde{C}_2$ that only depend on ρ and β/α but not on x, k, ℓ, h, or H.

Proof. Let $\zeta_x := 1 - \eta_x^{\ell(k-1)+1, \ell k - 1}$ with $\eta_x^{\ell(k-1)+1, \ell k - 1}$ as in equation (3.3) in Definition 3.3. Then there exists a $v \in V_h^{\mathrm{f}}(\omega_{x,\ell k})$ such that

$$
\begin{aligned}
|||\phi_x^h - v|||_{\omega_{x,\ell k}} &\leq |||\phi_x^h - \zeta_x \phi_x^h|||_{\omega_{x,\ell k}} + |||\zeta_x \phi_x^h - v|||_{\omega_{x,\ell k}} \\
&\lesssim |||\phi_x^h|||_{\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)+1}} + |||\zeta_x \phi_x^h|||_{\omega_{x,\ell k - 1} \setminus \omega_{x,\ell(k-1)+1}}.
\end{aligned}
$$

Furthermore, using the same argument as in Lemma 3.4,

$$|||\zeta_x \phi_x^h|||_{\omega_{x,\ell k - 1} \setminus \omega_{x,\ell(k-1)+1}} \lesssim |||\phi_x^h|||_{\omega_{x,\ell k} \setminus \omega_{x,\ell(k-1)+1}}$$

which yields

$$|||\phi_x^h - \phi_{x,\ell k}^h||| \leq |||\phi_x^h - v|||_{\omega_{x,\ell k}}^2 + |||\phi_x^h|||_{\Omega \setminus \omega_{x,\ell k}} \lesssim |||\phi_x^h|||_{\Omega \setminus \omega_{x,\ell(k-1)+1}}.$$

Now let $\eta_j := \eta_x^{\ell(j-1)+1,\ell j}$ (cf. Definition 3.3), $j = 2, 3, \ldots, k-1$ and note that

$$\|A^{1/2}\nabla\phi_x^h\|_{L^2(\Omega\setminus\omega_{x,\ell(k-1)})}^2 \leq a(\phi_x^h, \eta_{k-1}^2\phi_x^h) - 2\int_\Omega \eta_{k-1}\phi_x^h(A\nabla\phi_x^h)\cdot\nabla\eta_{k-1}.$$

The second term can be treated exactly as in Lemma 3.4 and, hence, bounded by $\ell^{-2} \left|\left|\left|\phi_x^h\right|\right|\right|_{\Omega\setminus\omega_{x,\ell(k-2)}}^2$. We make use of Lemma 4.7 to bound the first term. There exists $v \in V_h^{\mathrm{f}}(\Omega\setminus\omega_{x,\ell(k-1)+1})$ such that

$$a(\phi_x^h, \eta_{k-1}^2\phi_x^h) \leq \left|\left|\left|\phi_x^h\right|\right|\right|_{\omega_{\ell(k-1)}\setminus\omega_{\ell(k-2)+1}} \left|\left|\left|\eta_{k-1}^2\phi_x^h - v\right|\right|\right| \lesssim \ell^{-2} \left|\left|\left|\phi_x^h\right|\right|\right|_{\Omega\setminus\omega_{x,\ell(k-2)}}^2.$$

The final assertion follows by similar arguments as in the proof of Lemma 3.4. $\square$

Lemma 4.9 (Discrete version of Lemma 3.5). *There is a constant $\tilde{C}_3$ depending only on ρ and β/α, but not on $|\mathcal{N}|$, k, or ℓ such that*

$$\left|\left|\left|\sum_{x\in\mathcal{N}} v(x)(\phi_x^h - \phi_{x,\ell k}^h)\right|\right|\right|^2 \leq \tilde{C}_3(\ell k)^d \sum_{x\in\mathcal{N}} v^2(x)\left|\left|\left|\phi_x^h - \phi_{x,\ell k}^h\right|\right|\right|^2.$$

Proof. For $x \in \mathcal{N}$, let $\zeta_x = 1 - \eta_x^{\ell k+1,\ell k+2}$ (cf. Definition 3.3) and let $z = \sum_{x\in\mathcal{N}} v(x)(\phi_x - \phi_{x,\ell k})$. We have,

$$\left|\left|\left|\sum_{x\in\mathcal{N}} v(x)(\phi_x - \phi_{x,\ell k})\right|\right|\right|^2 = \sum_{x\in\mathcal{N}} v(x)\,a(\phi_x^h - \phi_{x,\ell k}^h, \zeta_x z + (1-\zeta_x)z) = \mathrm{I} + \mathrm{II}.$$

The first term $\mathrm{I} := \sum_{x\in\mathcal{N}} v(x)\,a(\phi_x^h - \phi_{x,\ell k}^h, \zeta_x z)$ can be treated in exactly the same way as in the proof of Lemma 3.5. We focus on the second term. Due to Lemma 4.7 there exists a $w \in V_h^{\mathrm{f}}(\Omega\setminus\omega_{\ell k})$ such that

$$\mathrm{II} := \sum_{x\in\mathcal{N}} v(x)\,a(\phi_x^h - \phi_{x,\ell k}^h, (1-\zeta_x)z - w)$$

$$\lesssim \left(\sum_{x\in\mathcal{N}} |v(x)|^2 \left|\left|\left|\phi_x^h - \phi_{x,\ell k}^h\right|\right|\right|^2\right)^{1/2} \left(\sum_{x\in\mathcal{N}} \left|\left|\left|(1-\zeta_x)z - w\right|\right|\right|^2\right)^{1/2}$$

$$\lesssim \left(\sum_{x\in\mathcal{N}} |v(x)|^2 \left|\left|\left|\phi_x^h - \phi_{x,\ell k}^h\right|\right|\right|^2\right)^{1/2} \left(\sum_{x\in\mathcal{N}} \left|\left|\left|z\right|\right|\right|_{\omega_{x,\ell k+2}\setminus\omega_{\ell k+1}}^2\right)^{1/2}$$

$$\lesssim (\ell k)^{d/2} \left(\sum_{x\in\mathcal{N}} |v(x)|^2 \left|\left|\left|\phi_x^h - \phi_{x,\ell k}^h\right|\right|\right|^2\right)^{1/2} \left|\left|\left|z\right|\right|\right|.$$

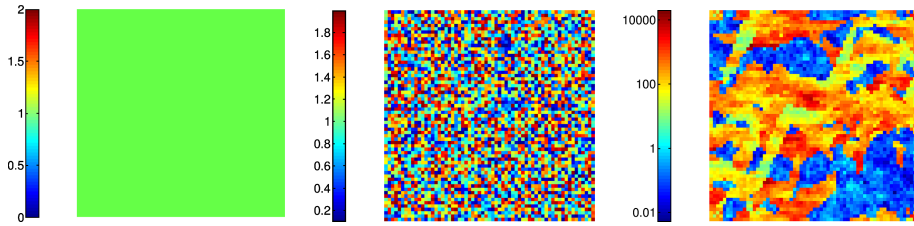The result follows immediately. $\square$

FIGURE 1. Scalar coefficient used in the numerical experiment: A_1 (left), A_2 (middle), A_3 (right).
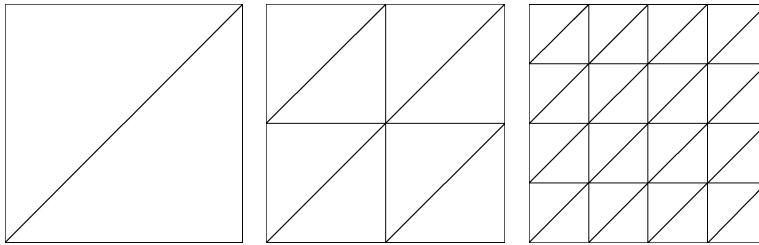

FIGURE 2. Uniform triangulations of the unit square.

5. NUMERICAL EXPERIMENTS

Numerical experiments shall validate our theoretical results from the previous sections.

5.1. **Experimental setup.** Let Ω be the unit square and the outer force $g \equiv 1$ in Ω. Consider three different choices for the scalar coefficient A_1, A_2, A_3 with increasing difficulty as depicted in Figure 1. The coefficient $A_1 = 1$ is constant. The coefficient A_2 is piecewise constant with respect to a uniform Cartesian grid of width 2^{-6}. The values in each grid cell are chosen in the range $[1/20, 2]$; the contrast $\beta(A_2)/\alpha(A_2) \le 40$ is moderate. The coefficient A_3 is piecewise constant with respect to the same uniform Cartesian grid of width 2^{-6}. Its values are taken from the data of the SPE10 benchmark; see http://www.spe.org/web/csp/. The contrast for A_3 is large, $\beta(A_3)/\alpha(A_3) \approx 4 \cdot 10^6$. Consider uniform coarse meshes of size $H = 2^{-1}, 2^{-2}, \ldots, 2^{-6}$ of Ω as depicted in Figure 2. Note that none of these meshes resolves the rough coefficients A_2 and A_3 appropriately.

The reference mesh $\mathcal{T}_h$ has width $h = 2^{-9}$. Since no analytical solutions are available, the standard finite element approximation $u_h \in V_h$ on the reference mesh $\mathcal{T}_h$ serves as the reference solution. All fine scale computations are performed on subsets of $\mathcal{T}_h$.

The approximations are compared with this reference solution only. Doing this, we assume that u_h is sufficiently accurate. True errors would behave similar in the beginning but level off at some point when the reference error $|||u - u_h|||$ dominates the upscaling error.

5.2. **Results for the energy error.** Figure 3 depicts the energy errors of the new multiscale method and the classical P1FEM (see (2.2)) with respect to the same coarse mesh. Depending on the coarse discretization scale H, the localization
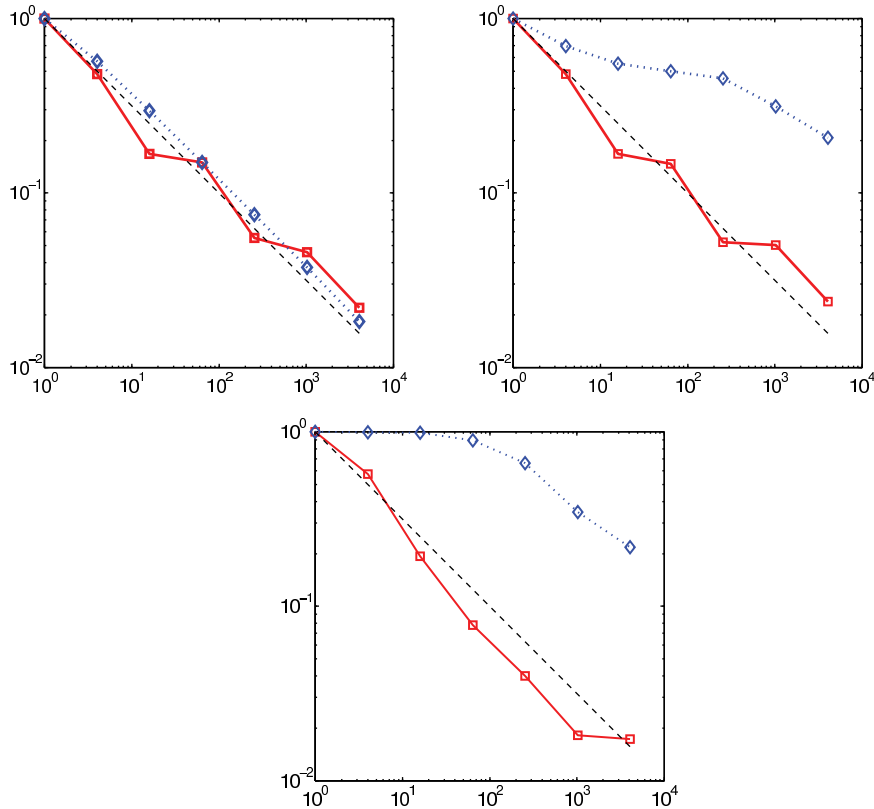
FIGURE 3. Relative energy errors $|||u_h - u_{H,k}^{\mathrm{ms},h}|||/|||u_h|||$ ($\square$ solid) with localization parameter $k = \lceil 2\log(1/H)\rceil$ and $|||u_h - u_H|||/|||u_h|||$ ($\diamond$ dotted) vs. number of degrees of freedom $N_{\mathrm{dof}} \approx H^{-2}$ for different coefficients: A_1 (top left), A_2 (top right), A_3 (bottom). The dashed black line is $N_{\mathrm{dof}}^{-1/2}$.

parameter k is chosen to be $\lceil 2\log(1/H)\rceil$. The logarithmic dependence on $1/H$ is motivated by our a priori analysis. The choice of the constant 2 is based on numerical tests. It turns out that, in all experiments, this choice leads to the desired linear textbook convergence (rate $-1/2$) of the energy error (w.r.t. to the number of degrees of freedom $N_{\mathrm{dof}} = |\mathcal{N}| \approx H^{-2}$) related to the sequence of multiscale approximations. Pre-asymptotic effects are not observed. In particular, the performance of our method does not seem to be affected by the high contrast present in A_3. Whether our estimates on the decay of the corrector functions are sub-optimal or have worst-case character with respect to contrast is an issue of present research.

Observe that the classical P1FEM suffers from the lacks of approximability and regularity and converges only poorly for the rough coefficients A_2 and A_3.

5.3. **Results for the L^2 error.** Figure 4 shows L^2 errors of the new multiscale method and the classical P1FEM. Again, the choice of the localization parameter $k = \lceil 2\log(1/H)\rceil$ yields the optimal convergence rate -1 for our method in all
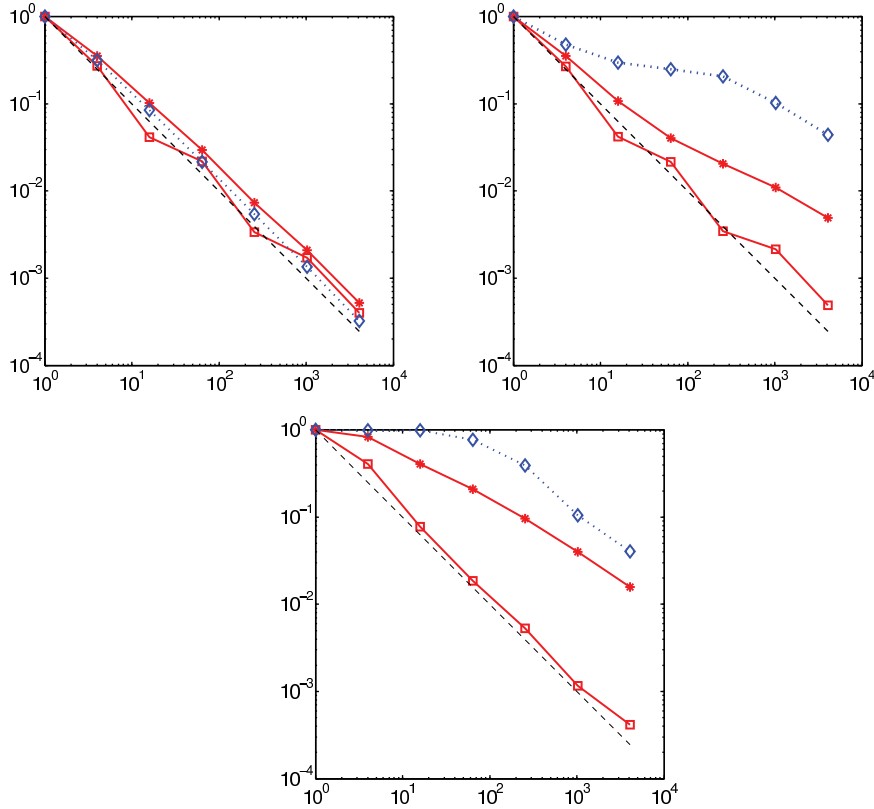
FIGURE 4. Relative L^2 errors $\|u_h - u_{H,k}^{\mathrm{ms},h}\|/\|u_h\|$ ($\square$ solid), $\|u_h - \mathfrak{I}_H u_{H,k}^{\mathrm{ms},h}\|/\|u_h\|$ ($*$ solid) with localization parameter $k = \lceil 2\log(1/H) \rceil$ and $\|u_h - u_H\|/\|u_h\|$ ($\Diamond$ dotted) vs. number of degrees of freedom $N_{\mathrm{dof}} \approx H^{-2}$ for different coefficients: A_1 (top left), A_2 (top right), A_3 (bottom). The dashed black line is N_{dof}^{-1}.

experiments (w.r.t. to the number of degrees of freedom $N_{\mathrm{dof}} = |\mathcal{N}| \approx H^{-2}$) without any pre-asymptotic behavior. This observation is justified by a standard Aubin-Nitsche duality argument. Define $e := u_h - u_{H,k}^{\mathrm{ms},h} \in L^2(\Omega)$ and let $z_h \in V_h$ solve

$$a(z_h, v_h) = \int_\Omega e v_h \quad \text{for all } v_h \in V_h.$$

Galerkin orthogonality leads to

$$\|u_h - u_{H,k}^{\mathrm{ms},h}\|_{L^2(\Omega)}^2 = a(z_e - z_{H,k}^{\mathrm{ms},h}, e) \leq \left|\left|\left| z_h - z_{H,k}^{\mathrm{ms},h} \right|\right|\right| \left|\left|\left| u_h - u_{H,k}^{\mathrm{ms},h} \right|\right|\right|,$$

where $z_{H,k}^{\mathrm{ms},h} \in V_{H,k}^{\mathrm{ms},h}$ is the Galerkin projection of z_h onto the discrete multiscale finite element space $V_{H,k}^{\mathrm{ms},h}$. The estimates for the energy error (see Section 4) and the present choice of k yield the L^2 estimate

$$\|u_h - u_{H,k}^{\mathrm{ms},h}\|_{L^2(\Omega)} \lesssim H^2 \|g\|_{L^2(\Omega)}.$$

More importantly, we observe that the L^2 error between u_h and $\mathfrak{I}_H u_{H,k}^{\mathrm{ms},h}$ converges nicely at a rate close to $-3/4$ without pre-asymptotic effects. This is remarkable because $\mathfrak{I}_H u_{H,k}^{\mathrm{ms},h}$ is a truly coarse approximation. $\mathfrak{I}_H u_{H,k}^{\mathrm{ms},h}$ is an element of the coarse $P1$ finite element space. Hence, it cannot capture microscopic features of the solution. The rate of convergence (with respect to the number of degrees of freedom) is limited by $\frac{1+s}{2}$ for some $s \in [0,1]$ which is related to the regularity of the solution ($u \in H^{1+s}$ for some $s \in [0,1]$). However, $\mathfrak{I}_H u_{H,k}^{\mathrm{ms},h}$ approximates the macroscopic behavior of the solution accurately with only very few degrees of freedom. Note that the storage complexity of the modified basis is of order $\mathcal{O}(h^{-2}\log 1/H)$ whereas its interpolation can be stored in $\mathcal{O}(H^{-2}\log 1/H)$. Once the coarse system matrix of the multiscale method is assembled, $\mathfrak{I}_H u_{H,k}^{\mathrm{ms},h}$ can be computed without using any fine scale information from the modified basis whereas this would be required to represent the full multiscale approximation $u_{H,k}^{\mathrm{ms},h}$.

6. Application to multiscale methods

In this section we discuss three multiscale methods and how the presented analysis relates to each of them.

6.1. The variational multiscale method.

The variational multiscale method was first introduced in [14]. The function space V is here split into a coarse part (standard finite element space on a coarse mesh), in our case V_H, and a fine part, in our case V^{f}. The weak form is also decoupled into a coarse and a fine part. The method reads: find $\bar{u} \in V_H$ and $u' \in V^{\mathrm{f}}$ such that

$$a(\bar{u}, \bar{v}) + a(u', \bar{v}) = G(\bar{v}) \quad \text{for all } \bar{v} \in V_H,$$
$$a(u', v') = G(v') - a(\bar{u}, v') \quad \text{for all } v' \in V^{\mathrm{f}}.$$

The fine scale solution is further decoupled over the coarse elements $T \in \mathcal{T}_H$ and approximated using analytical techniques. Note that the fine scale solution u' is an affine map of the coarse scale solution $\bar{u}$. If we let $u' \approx M\bar{u} + m$ and plug this in to the first equation we get a coarse stiffness matrix of the form $a(\bar{v} + M\bar{v}, \bar{w})$, i.e., a non-symmetric bilinear form for a symmetric problem.

6.2. The multiscale finite element method.

In [13] the multiscale finite element method was first introduced. Here modified multiscale basis functions are computed numerically on sub-grids on each coarse element individually. The corrector functions fulfill: find $\phi_{x,T} \in H_0^1(T)$,

$$a(\lambda_x - \phi_{x,T}, v) = 0 \quad \text{for all } v \in H_0^1(T) \text{ and for all } T \in \mathcal{T}_H.$$

Here homogeneous Dirichlet boundary conditions are used on the boundary of each element T, i.e., the local problems are totally decoupled. To get a more accurate method one can improve the boundary conditions using information from the data A. A larger domain can also be considered (this procedure is referred to as oversampling); see [13]. Note that since the coarse scale basis functions are modified (both trial and test space) the resulting method is symmetric.

6.3. **The adaptive variational multiscale method.** The modified basis function construction given by equation (2.7) and (2.8) was first introduced in a variational multiscale framework in [15, 16]. In these papers the Scott-Zhang interpolation was used in the analysis and nodal interpolation in the discrete setting for the numerical examples. The modified basis functions where only used for the trial functions but not for the test functions. A fine scale correction based on the right-hand side data was also included. In [18] the modified basis functions were used for both trial and test functions. The exponential decay of the modified basis functions, with respect to the number of coarse layers of elements in the vertex patches, has been demonstrated numerically in all these works; see [17, 18].

The adaptive variational multiscale method has been extended to convection dominated problems and problems in mixed form [18]. A posteriori error bounds have been derived and adaptive algorithms designed where the local mesh and patch size are chosen automatically in order to reduce the error.

6.4. **Application of the presented analysis.** The convergence proof in this paper gives a valid bound also as $h \to 0$ independent of the patch size and coarse mesh size. The proof does not rely on regularity of the solution and gives a very explicit expression for the rate of convergence. The present analysis confirms the numerical results in [17, 18] and gives the symmetric version of the method, where both trial and test space are modified, the solid theoretical foundation it has previously been missing. The analysis also justifies the use of a posteriori error bounds for adaptivity [16, 18] because we can now prove that the quantities measured on the patch boundary decays exponentially in the number of coarse layers.

For the variational multiscale method this result says that it is important to allow larger subgrid patches than just one coarse element. This will result in overlap but the local problems are totally decoupled and we have in previous works demonstrated how adaptivity can be used to only solve local problems where it is needed, see for instance [16, 18]. For the multiscale finite element method the analysis is not directly applicable since the fine scale space V^{f} is not used. It is the decay in this space which we have proven to be exponential (in number of coarse layers of elements in the subgrid). If this decay is not present, inhomogeneous boundary conditions are instead needed for the subgrid problems. To the best of our knowledge, such constructions have only been proved to be accurate in special settings, e.g., periodic coefficients.

References

[1] I. Babuška and J. E. Osborn, *Generalized finite element methods: their performance and their relation to mixed methods*, SIAM J. Numer. Anal. **20** (1983), no. 3, 510–536, DOI 10.1137/0720034. MR701094 (84h:65076)

[2] Ivo Babuška, Gabriel Caloz, and John E. Osborn, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, SIAM J. Numer. Anal. **31** (1994), no. 4, 945–981, DOI 10.1137/0731051. MR1286212 (95g:65146)

[3] Ivo Babuska and Robert Lipton, *Optimal local approximation spaces for generalized finite element methods with application to multiscale problems*, Multiscale Model. Simul. **9** (2011), no. 1, 373–406, DOI 10.1137/100791051. MR2801210 (2012e:65259)

[4] Ivo Babuška and John E. Osborn, *Can a finite element method perform arbitrarily badly?*, Math. Comp. **69** (2000), no. 230, 443–462, DOI 10.1090/S0025-5718-99-01085-6. MR1648351 (2000i:65114)

[5] I. Babuška and J. M. Melenk, *The partition of unity method*, Internat. J. Numer. Methods Engrg. **40** (1997), no. 4, 727–758, DOI 10.1002/(SICI)1097-0207(19970228)40:4⟨727::AID-NME86⟩3.3.CO;2-E. MR1429534 (97j:73071)

[6] Leonid Berlyand and Houman Owhadi, *Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast*, Arch. Ration. Mech. Anal. **198** (2010), no. 2, 677–721, DOI 10.1007/s00205-010-0302-1. MR2721592 (2012b:35016)

[7] F. Brezzi, L. P. Franca, T. J. R. Hughes, and A. Russo, $b = \int g$, Comput. Methods Appl. Mech. Engrg. **145** (1997), no. 3-4, 329–339, DOI 10.1016/S0045-7825(96)01221-2. MR1456019 (98g:65086)

[8] Carsten Carstensen, *Quasi-interpolation and a posteriori error analysis in finite element methods*, M2AN Math. Model. Numer. Anal. **33** (1999), no. 6, 1187–1202, DOI 10.1051/m2an:1999140. MR1736895 (2001a:65135)

[9] Carsten Carstensen and Rüdiger Verfürth, *Edge residuals dominate a posteriori error estimates for low order finite element methods*, SIAM J. Numer. Anal. **36** (1999), no. 5, 1571–1587 (electronic), DOI 10.1137/S003614299732334X. MR1706735 (2000g:65115)

[10] Philippe G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4. MR0520174 (58 #25001)

[11] Ph. Clément, *Approximation by finite element functions using local regularization* (English, with Loose French summary), Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique **9** (1975), no. R-2, 77–84. MR0400739 (53 #4569)

[12] L. Grasedyck, I. Greff, and S. Sauter, *The AL basis for the solution of elliptic problems in heterogeneous media*, Multiscale Model. Simul. **10** (2012), no. 1, 245–258, DOI 10.1137/11082138X. MR2902606

[13] Thomas Y. Hou and Xiao-Hui Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys. **134** (1997), no. 1, 169–189, DOI 10.1006/jcph.1997.5682. MR1455261 (98e:73132)

[14] Thomas J. R. Hughes, Gonzalo R. Feijóo, Luca Mazzei, and Jean-Baptiste Quincy, *The variational multiscale method—a paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg. **166** (1998), no. 1-2, 3–24, DOI 10.1016/S0045-7825(98)00079-6. MR1660141 (99m:65239)

[15] Mats G. Larson and Axel Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation: duality techniques for elliptic problems*, Multiscale methods in science and engineering, Lect. Notes Comput. Sci. Eng., vol. 44, Springer, Berlin, 2005, pp. 181–193, DOI 10.1007/3-540-26444-2_9. MR2161713 (2006f:65119)

[16] Mats G. Larson and Axel Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation: energy norm estimates for elliptic problems*, Comput. Methods Appl. Mech. Engrg. **196** (2007), no. 21-24, 2313–2324, DOI 10.1016/j.cma.2006.08.019. MR2319044 (2008e:65328)

[17] Mats G. Larson and Axel Målqvist, *A mixed adaptive variational multiscale method with applications in oil reservoir simulation*, Math. Models Methods Appl. Sci. **19** (2009), no. 7, 1017–1042, DOI 10.1142/S021820250900370X. MR2553176 (2010i:65274)

[18] Axel Målqvist, *Multiscale methods for elliptic problems*, Multiscale Model. Simul. **9** (2011), no. 3, 1064–1086, DOI 10.1137/090775592. MR2831590 (2012j:65419)

[19] Houman Owhadi and Lei Zhang, *Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast*, Multiscale Model. Simul. **9** (2011), no. 4, 1373–1398, DOI 10.1137/100813968. MR2861243 (2012k:35037)

[20] D. Peterseim and S. Sauter, *Finite elements for elliptic problems with highly varying, nonperiodic diffusion matrix*, Multiscale Model. Simul. **10** (2012), no. 3, 665–695, DOI 10.1137/10081839X. MR3022017

[21] Harry Yserentant, *On the multilevel splitting of finite element spaces*, Numer. Math. **49** (1986), no. 4, 379–412, DOI 10.1007/BF01389538. MR853662 (88d:65068a)

DEPARTMENT OF MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY AND UNIVERSITY OF GOTHENBURG, CHALMERS TVÄRGATA 3, SE-14296 GÖTEBORG, SWEDEN
  E-mail address: `axel@chalmers.se`

RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN, INSTITUTE FOR NUMERICAL SIMULATION, WEGELERSTR. 6, 53115 BONN, GERMANY
  E-mail address: `peterseim@ins.uni-bonn.de`