Localization of Elliptic Multiscale Problems

Axel Målqvist [*] Daniel Peterseim [†‡]

October 4, 2011

Abstract

This note constructs a local generalized finite element basis for elliptic problems with heterogeneous and highly varying diffusion tensor. The basis functions are solutions of local problems on vertex patches. The error of the corresponding generalized finite element method decays exponentially w.r.t. the number of element layers in the patches. Hence, on a uniform mesh of size H, patches of diameter $\log(1/H)$ are sufficient to preserve the convergence rates of the classical P_1-FEM for the Poisson problem. The analysis does not rely on regularity of the solution or scale separation in the coefficient. The result justifies the use of the class of variational multiscale methods, introduced in [Comput. Methods Appl. Mech. Engrg., 196:2313–2324, 2007].

Keywords finite element method, a priori error estimate, convergence, multiscale method

AMS subject classifications 65N12, 65N30

1 Introduction

It is well known that classical polynomial based finite element methods might perform arbitrarily badly for elliptic problems with strongly heterogeneous and varying diffusion coefficient, see e.g. [4]. In the present context, the heterogeneities

[*]Department of Information Technology, Uppsala University, Box 337, SE-751 05 Uppsala, Sweden

[†]Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

[‡]The second author is supported by the DFG Research Center Matheon Berlin.

and oscillations of the coefficient might appear on several, possibly non-separated, scales. To overcome the lack of performance in such cases, many methods that are based on general (non-polynomial) ansatz functions have been developed. Early work [3, 1], that applies to essentially one dimensional problems has been generalized to the multi-dimensional case in several ways during the last fifteen years, see e.g. [12, 11]. In these methods the problem is split into a coarse and a (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 very promising results in practice, their convergence analysis usually assume certain periodicity and scale separation of the coefficient.

For problems with general L^∞ coefficient, [2] gives error bounds for a generalized finite element method that involves the solutions of local eigenvalue problems. The construction [6] only depends on the solution of the original problem on certain subdomains. The size of these subdomains strongly depends on the mesh size. This is suboptimal with regard to the theoretical statement given in [10], that is, for any shape regular mesh of size H there exist $O\left((\log(1/H))^{d+1}\right)$ local (non-polynomial) basis functions per nodal point such that the error of the corresponding Galerkin solution u_H satisfies the estimate $\|u - u_H\|_{H^1(\Omega)} \leq C_f H$ with a constant C_f that depends on f and the global bounds of the diffusion coefficient but not on its variations. However, the derivation 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 basis can indeed be constructed by solving only local problems on element patches. We use the modified nodal basis presented in [14] and prove that these basis functions decay exponentially away from the node they are associated with. This exponential decay justifies an approximation on localized patches.

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

$$\begin{aligned} 0 < \alpha(A,\Omega) &:= \operatorname*{ess\,inf}_{x\in\Omega} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\langle A(x)v, v\rangle}{\langle v, v\rangle}, \\ \infty > \beta(A,\Omega) &:= \operatorname*{ess\,sup}_{x\in\Omega} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\langle A(x)v, v\rangle}{\langle v, v\rangle}. \end{aligned} \tag{1}$$

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

$$a(u,v) := \int_\Omega \langle A\nabla u, \nabla v\rangle \,\mathrm{d}x = \int_\Omega fv \,\mathrm{d}x =: F(v) \quad \text{for all } v \in V. \tag{2}$$

The bilinear form a is symmetric, coercive, bounded, and hence, (2) has a unique solution.

The main result of this paper shows that the error $u - u_H^{\text{ms}}$ of the generalized finite element method, which is based on our modified (local) basis functions mentioned above, is bounded as follows (cf. Theorem 5)

$$\|A^{1/2}\nabla(u - u_H^{\text{ms}})\| \leq C_f(H + \gamma^k);$$

H being the mesh size of the underlying coarse finite element mesh, k referring to the number of element layers that form the support of the of the localized basis functions, and $\gamma < 1$. Moreover, this result 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 the exponential decay of the modified basis has been noticed, see e.g. [17].

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 and Section 5 discusses application of this theory to several state-of-the-art multiscale finite element methods.

2 Local Basis

This section designs a set of local basis functions for the multiscale problem under consideration. The construction is based on a regular (in the sense of [8]) finite element mesh $\mathcal{T}$ of Ω into closed triangles ($d = 2$) or tetrahedra ($d = 3$). Subsection 2.1 recalls the classical nodal basis with respect to $\mathcal{T}$ and motivates its lack of approximation properties. Subsection 2.2 defines a modified (coefficient dependent) nodal basis and analyzes its approximation properties. This basis is localized in Subsection 2.3.

2.1 Classical Nodal Basis

Let $H : \overline{\Omega} \to \mathbb{R}_{>0}$ denote the $\mathcal{T}$-piecewise constant mesh size function with $H|_T = \operatorname{diam}(T)$ for all $T \in \mathcal{T}$. The mesh size might vary in space. In practical applications, the mesh $\mathcal{T}$ (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

$$S_H := \{v \in C^0(\Omega) \mid \forall T \in \mathcal{T}, v|_T \text{ is a polynomial of total degree} \leq 1\}. \tag{3.a}$$

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

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

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 linear system is sparse.

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

$$a(u_H, v) = F(v) \quad \text{for all } v \in V_H. \tag{3.b}$$

The above method (3) 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. It holds

$$|||u - u_H||| = \min_{v_H \in V_H} |||u - v_H|||. \tag{4}$$

Assuming that the solution u is smooth, the combination of (4) 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 (3) 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 might oscillate rapidly, say at frequency ε^{-1} for some small parameter ε. In such a situation the asymptotic result is useless, because $\nabla^2 u$ might oscillate at the same scale, a fact, which 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 modification of this method that resolves this issue.

2.2 Multiscale Splitting and Modified Nodal Basis

Let $\mathfrak{I}_\mathcal{T} : V \to V_H$ denote the Clément interpolation operator [9, 19]. Then the space

$$V^{\mathrm{f}} := \{v \in V \mid \mathfrak{I}_\mathcal{T} v = 0\}$$

represents the microscopic features of V that are not captured by V_H. For $v \in V_H$ define $\mathfrak{F}v$ 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

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

Hence, every function $u \in V$ can be split 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 corresponding Galerkin approximation $u_H^{\mathrm{ms}} \in V_H^{\mathrm{ms}}$ satisfies

$$a(u_H^{\mathrm{ms}}, v) = F(v) \quad \text{for all } v \in V_H^{\mathrm{ms}}. \tag{5}$$

The error $(u - u_H^{\mathrm{ms}})$ of the above method (5) is analyzed in Section 3.1.

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

$$a(\phi_x, w) = a(\lambda_x, w) \quad \text{for all } w \in V^{\mathrm{f}}. \tag{6}$$

Hence, a basis of V_H^{ms} is given by the modified nodal basis

$$\{\lambda_x - \phi_x \mid x \in \mathcal{N} \cap \Omega\}. \tag{7}$$

In general, the corrections ϕ_x of nodal basis functions λ_x, $x \in \mathcal{N}$ have global support. Hence, the modified basis (7) and the corresponding method (5) is of limited use in practice.

2.3 Localization

In this section, we will show that the correction ϕ_x decays exponentially fast away from x and that a simple truncation leads to localized basis functions with good approximation properties.

Let $k \in \mathbb{N}$. Define nodal patches of k-th order ω_x^k around $x \in \mathcal{N}$ by

$$\begin{aligned} \omega_{x,1} &:= \operatorname{int}\left(\cup\{T \in \mathcal{T} \mid x \in T\}\right), \\ \omega_{x,k} &:= \operatorname{int}\left(\cup\{T \in \mathcal{T} \mid T \cap \overline{\omega}_{x,k-1} \neq \emptyset\}\right), \quad k = 2, 3, 4 \ldots. \end{aligned} \tag{8}$$

Then, the solutions $\phi_{x,k} \in V^{\mathrm{f}}(\omega_{x,k}) := V^{\mathrm{f}} \cap H_0^1(\omega_{x,k})$ of

$$a(\phi_{x,k}, w) = a(\lambda_x, w) \quad \text{for all } w \in V^{\mathrm{f}}(\omega_{x,k}), \tag{9}$$

are approximations of ϕ_x from (6) which have local supports. By extending $\phi_{x,k}$ by zero in $\Omega \setminus \omega_{x,k}$ we have that $\phi_{x,k} \in V^{\mathrm{f}}$.

We define localized multiscale finite element spaces

$$V_{H,k}^{\mathrm{ms}} = \operatorname{span}\{\lambda_x - \phi_{x,k} \mid x \in \mathcal{N}\}. \tag{10.a}$$

The corresponding multiscale approximation of (2) reads: Find $u_{H,k}^{\mathrm{ms}} \in V_{H,k}^{\mathrm{ms}}$ such that

$$a(u_{H,k}^{\mathrm{ms}}, v) = F(v) \quad \text{for all } v \in V_{H,k}^{\mathrm{ms}}. \tag{10.b}$$

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 (10) is the same as for the classical method (3). 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 1 *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. This procedure leads to $O\left((\log(1/H))^d\right)$ local basis functions per vertex which span a generalized finite element space with similar approximation properties as $V_{H,k}^{ms}$ [5].*

3 Error Analysis

This section analyzes the proposed multiscale method in three steps. First, Subsection 3.1 presents an error bound for the idealized method (5). 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 (10).

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

$$\rho := \max_{T \in \mathcal{T}} \rho_T \quad \text{with} \quad \rho_T := \frac{\operatorname{diam} B_T}{\operatorname{diam} T} \ \text{ for } T \in \mathcal{T}, \tag{11}$$

where B_T denotes the largest ball contained in T. In addition, a moderate dependence on the ratio $C_H := \|H\|_{L^\infty(\Omega)} \|H^{-1}\|_{L^\infty(\Omega)}$ appears. However, its influence on the total error can be controlled by the localization parameter k.

3.1 Discretization Error

Lemma 2 *Let $u \in V$ solve* (2) *and $u_H^{\mathrm{ms}} \in V_H^{\mathrm{ms}}$ solve* (5)*. Then it holds*

$$|||u - u_H^{\mathrm{ms}}||| \leq C_{\mathfrak{I}_{\mathcal{T}}} \alpha^{-1} \|Hf\|_{L^2(\Omega)}.$$

Proof. Recall the (local) approximation and stability properties of the interpolation operator $\mathfrak{I}_{\mathcal{T}}$ [20, 7]: There exists a generic (computable) constant $C_{\mathfrak{I}_{\mathcal{T}}}$ which only depends on ρ_T but not on $\operatorname{diam} T$ such that for all $k, m \in \mathbb{N}$ with $0 \leq k \leq m$ and for all $v \in H^m(\Omega)$ and for all $T \in \mathcal{T}$ it holds

$$\|\nabla^k (v - \mathfrak{I}_{\mathcal{T}} v)\|_{L^2(T)} \leq C_{\mathfrak{I}_{\mathcal{T}}} (\operatorname{diam} T)^{m-k} \|\nabla^m v\|_{L^2(\omega_T)} \tag{12}$$

with $\omega_T := \cup\{K \in \mathcal{T} \mid T \cap K \neq \emptyset\}$.

Due to the splitting from Section 2.2, it holds $u - u_H^{\mathrm{ms}} = u^{\mathrm{f}}$. Since $\mathfrak{I}_{\mathcal{T}} u^{\mathrm{f}} = 0$, the estimate (12) and the finite overlap of the patches ω_T conclude the proof,

$$|||u^{\mathrm{f}}|||^2 = F(u^{\mathrm{f}}) \leq \sum_{T\in\mathcal{T}} \|f\|_{L^2(T)} \|u^{\mathrm{f}} - \mathfrak{I}_{\mathcal{T}} u^{\mathrm{f}}\|_{L^2(T)} \leq C_{\mathfrak{I}_{\mathcal{T}}} \alpha^{-1} \|Hf\|_{L^2(\Omega)} |||u^{\mathrm{f}}|||.$$

■

3.2 Error of Localized Multiscale FEM

First, we estimate the error due to truncation to local patches.

Lemma 3 *For all $x \in \mathcal{N}$, $k \geq 2 \in \mathbb{N}$, and $\ell \in \mathbb{N}_{>1}$, the estimate*

$$|||\phi_x - \phi_{x,\ell k}||| \leq C_2 \left(\frac{C_1}{\ell}\right)^{\frac{k-1}{2}} |||\phi_x|||_{\omega_{x,\ell}}$$

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

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

$$|||\phi_x - \phi_{x,\ell k}|||^2 = |||\phi_x - \phi_{x,\ell k}|||^2_{\omega_{x,\ell k}} + |||\phi_x|||^2_{\Omega\setminus\omega_{x,\ell k}}.$$

The function $\phi_x - \phi_{x,\ell k}$ is an element of the linear space $X := \{\phi \in V^{\mathrm{f}} \mid \forall \psi \in V^{\mathrm{f}}(\omega_{x,\ell k}),\ a(\phi,\psi) = 0\}$; its trace $(\phi_x - \phi_{x,\ell k})|_{\omega_{x,\ell k}}$ belongs to the linear space $Y := \{g \in L^2(\partial\omega_{x,\ell k}) \mid \exists \phi \in X :\ \phi|_{\partial\omega_{x,\ell k}} = g \text{ in the sense of traces}\}$. The trace operator $\cdot|_{\partial\omega_{x,\ell k}}$ is a continuous linear mapping from the Banach space $(X, |||\cdot|||_{\omega_{x,\ell k}})$ onto the Banach space $(Y, \|\cdot\|_{L^2(\partial\omega_{x,\ell k})})$. Given any $\phi \in X$, observe that $\|\phi|_{\partial\omega_{x,\ell k}}\|_{L^2(\partial\omega_{x,\ell k})} = 0$ implies $\|\phi|_{\partial\omega_{x,\ell k}}\|_{H^{\frac{1}{2}}(\partial\omega_{x,\ell k})} = 0$ and, hence, $|||\phi|||_{\omega_{x,\ell k}} = 0$. Therefore, the trace operator is a bijection between X and Y and, by the inverse mapping theorem, its inverse is bounded, i.e., there is a constant C_2' such that

$$|||\phi|||_{\omega_{x,\ell k}} \leq C_2' \operatorname{diam}(\omega_{x,\ell k})^{-\frac{1}{2}} \|\phi\|_{L^2(\partial\omega_{x,\ell k})} \qquad \text{for all } \phi \in X.$$

A scaling argument shows that C_2' does not depend on $\operatorname{diam}(\omega_{x,\ell k})$. Since $\phi_{x,\ell k}|_{\partial\omega_{x,\ell k}} = 0$, this yields

$$|||\phi_x - \phi_{x,\ell k}|||_{\omega_{x,\ell k}} \leq C_2' \operatorname{diam}(\omega_{x,\ell k})^{-\frac{1}{2}} \|\phi_x\|_{L^2(\partial\omega_{x,\ell k})}.$$

The trace theorem for H^1-functions on simplices [7, Remark 4.1], together with $\mathfrak{I}_{\mathcal{T}}\phi_x = 0$ and (12), shows that

$$C_2' \operatorname{diam}(\omega_{x,\ell k})^{-\frac{1}{2}} \|\phi_x\|_{L^2(\partial\omega_{x,\ell k})} \leq C_2 |||\phi_x|||_{\Omega\setminus\omega_{x,\ell k}}$$

holds with a constant C_2 that does not depend on the diameter of $\omega_{x,\ell k}$. Thus,

$$|||\phi_x - \phi_{x,\ell k}||| \le C_2 |||\phi_x|||_{\Omega\setminus\omega_{x,\ell k}}. \tag{13}$$

Further estimation of the right hand side in (13) is possible using cutoff functions $\eta_j : \Omega \to [0,1] \in V$, $j = 2, 3, \dots, k$ with

$$(\eta_j)|_{\omega_{x,\ell(j-1)+1}} = 0, \tag{14.a}$$
$$(\eta_j)|_{\Omega\setminus\omega_{x,\ell j}} = 1, \tag{14.b}$$
$$\forall T \in \mathcal{T},\ \|\nabla\eta_j\|_{L^\infty(T)} \le C_\eta \ell^{-1} \operatorname{diam}(T)^{-1}. \tag{14.c}$$

For example, one might choose $\eta_j \in S_H$ with nodal values

$$\begin{aligned}
\eta_j(x) &= 0 \quad \text{for all } x \in \mathcal{N} \cap \omega_{x,\ell(j-1)+1},\\
\eta_j(x) &= 1 \quad \text{for all } x \in \mathcal{N} \cap \left(\Omega \setminus \omega_{x,\ell j}\right), \text{ and}\\
\eta_j(x) &= \frac{m}{\ell - 1} \quad \text{for all } x \in \mathcal{N} \cap \partial\omega_{x,\ell(j-1)+1+m},\ m = 0, 1, 2, \dots, \ell - 1.
\end{aligned}$$

With the above choice, (14.c) holds with a constant C_η that only depends on ρ. Since $|\operatorname{supp}(\nabla\lambda_x) \cap \operatorname{supp}(\eta_k)| = 0$ and $\operatorname{supp}(\nabla\eta_k) = \omega_{x,k\ell} \setminus \omega_{x,(k-1)\ell+1}$ we have

$$\begin{aligned}
\|A^{1/2}\nabla\phi_x\|^2_{L^2(\Omega\setminus\omega_{x,\ell k})} &\le \|A^{1/2}\eta_k\nabla\phi_x\|^2_{L^2(\Omega)}\\
&= \int_\Omega \left\langle A\nabla\phi_x, \nabla(\eta_k^2\phi_x)\right\rangle \mathrm{d}x \;-\; 2\int_\Omega \eta_k\phi_x \left\langle A\nabla\phi_x, \nabla\eta_k\right\rangle \mathrm{d}x\\
&= \int_\Omega \left\langle A\nabla\lambda_x, \nabla(\eta_k^2\phi_x)\right\rangle \mathrm{d}x - 2\int_\Omega \eta_k\phi_x \left\langle A\nabla\phi_x, \nabla\eta_k\right\rangle \mathrm{d}x\\
&\le 2 \sum_{T\in\mathcal{T}:\, T\subset\overline{\omega}_{x,k\ell}\setminus\omega_{x,(k-1)\ell+1}} \|\nabla\eta_k\|_{L^\infty(T)} \|A^{1/2}\nabla\phi_x\|_{L^2(T)} \|A^{1/2}\phi_x\|_{L^2(T)}.
\end{aligned}$$

The property (14.c) of the cutoff function and the upper bound of the interpolation error (12) yield

$$\|A^{1/2}\nabla\phi_x\|^2_{L^2(\Omega\setminus\omega_{x,\ell k})} \le C_1\ell^{-1}\|A^{1/2}\nabla\phi_x\|^2_{L^2(\Omega\setminus\omega_{x,(k-1)\ell})}, \tag{15}$$

where $C_1 := 2C_{\mathfrak{I}_\mathcal{T}}\sqrt{C_A}C_\eta$ with $C_A := \max_{T\in\mathcal{T}} \frac{\beta(A,T)}{\alpha(A,T)}$. For $j = k, k-1, \dots, 2$, a similar argument (with η_k replaced by η_j) yields

$$\|A^{1/2}\nabla\phi_x\|^2_{L^2(\Omega\setminus\omega_{x,(j-1)\ell})} \le C_1\ell^{-1}\|A^{1/2}\nabla\phi_x\|^2_{L^2(\Omega\setminus\omega_{x,(j-2)\ell})}. \tag{16}$$

Starting from (15), the successive application of (16) for $j = k, k-1, k-2, \dots, 2$ proves

$$\|A^{1/2}\nabla\phi_x\|^2_{L^2(\Omega\setminus\omega_{x,\ell k})} \le (C_1\ell^{-1})^{k-1}\|A^{1/2}\nabla\phi_x\|^2_{L^2(\omega_{x,\ell})}. \tag{17}$$

Combining (13) and (17), we finally obtain the assertion. ∎

Lemma 4 *There is a constant C_3 depending only on ρ, C_H, and β/α, but not on $|\mathcal{N}|$, k, or ℓ such that,*

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

Proof. There exists continuous functions $\zeta_x : \Omega \to [0,1]$ such that,

$$\begin{gathered}(\zeta_x)|_{\omega_{x,\ell l}} = 1, \quad (\zeta_x)|_{\Omega\setminus\omega_{x,\ell k+1}} = 0,\\ \|\nabla\zeta_x\|_{L^\infty(T)} \leq \frac{C'}{\operatorname{diam}(T)} \quad \text{for some constant } C', \text{ and}\\ \forall v \in V^{\mathrm{f}},\ \zeta_x v \in V^{\mathrm{f}}.\end{gathered}$$

We note that $a(\phi_x, (1-\zeta_x)v) = 0$ for all $v \in V^{\mathrm{f}}$. Furthermore also $a(\phi_{x,\ell k}, (1-\zeta_x)v) = 0$ for all $v \in V^{\mathrm{f}}$ since the functions support are disjoint. We get,

$$\begin{aligned}\left|\!\left|\!\left|\sum_{x\in\mathcal{N}} v(x)\,(\phi_x - \phi_{x,\ell k})\right|\!\right|\!\right|^2 &= \sum_{x\in\mathcal{N}} v(x)\, a(\phi_x - \phi_{x,\ell k}, \sum_{y\in\mathcal{N}} v(y)(\phi_y - \phi_{y,\ell k}))\\
&= \sum_{x\in\mathcal{N}} v(x)\, a(\phi_x - \phi_{x,\ell k}, \zeta_x \sum_{y\in\mathcal{N}} v(y)(\phi_y - \phi_{y,\ell k}))\\
&\leq \sum_{x\in\mathcal{N}} v(x)\, |||\phi_x - \phi_{x,\ell k}||| \cdot \left|\!\left|\!\left|\zeta_x \sum_{y\in\mathcal{N}} v(y)(\phi_y - \phi_{y,\ell k})\right|\!\right|\!\right|\\
&\leq \sum_{x\in\mathcal{N}} v(x)\, |||\phi_x - \phi_{x,\ell k}||| \cdot \beta^{1/2} \left\|(\nabla\zeta_x)(1-\mathfrak{I}_{\mathcal{T}}) \sum_{y\in\mathcal{N}} v(y)(\phi_y - \phi_{y,\ell k})\right\|_{L^2(\omega_{x,\ell k+1})}\\
&\quad + \sum_{x\in\mathcal{N}} v(x)\, |||\phi_x - \phi_{x,\ell k}||| \cdot \left|\!\left|\!\left|\sum_{y\in\mathcal{N}} v(y)(\phi_y - \phi_{y,\ell k})\right|\!\right|\!\right|_{\omega_{x,\ell k+1}}\\
&\leq C'' \sum_{x\in\mathcal{N}} v(x)\, |||\phi_x - \phi_{x,\ell k}||| \cdot \left|\!\left|\!\left|\sum_{y\in\mathcal{N}} v(y)(\phi_y - \phi_{y,\ell k})\right|\!\right|\!\right|_{\omega_{x,\ell k+2}}\\
&\leq C'' \left(\sum_{x\in\mathcal{N}} v^2(x)|||\phi_x - \phi_{x,\ell k}|||^2\right)^{1/2} \left(\sum_{x\in\mathcal{N}} \left|\!\left|\!\left|\sum_{y\in\mathcal{N}} v(y)(\phi_y - \phi_{y,\ell k})\right|\!\right|\!\right|^2_{\omega_{x,\ell k+2}}\right)^{1/2}\\
&\leq C_3(\ell k)^{d/2} \left(\sum_{x\in\mathcal{N}} v^2(x)|||\phi_x - \phi_{x,\ell k}|||^2\right)^{1/2} \left|\!\left|\!\left|\sum_{x\in\mathcal{N}} v(x)(\phi_x - \phi_{x,\ell k})\right|\!\right|\!\right|,\end{aligned}$$

where C_3 depends on C', C_H, $C_{\mathfrak{I}_{\mathcal{T}}}$, and β/α. The result follows by dividing by $|||\sum_{x\in\mathcal{N}} v(x)(\phi_x - \phi_{x,\ell k})|||$ on both sides. ∎

Theorem 5 *Let $u \in V$ solve* (2) *and, given $k \geq 2, \ell \in \mathbb{N}$, let $u_{H,\ell k} \in V^{ms}_{H,\ell k}$ solve* (10)*. Then*

$$|||u - u_{H,\ell k}||| \leq C_4 \left(\ell \sqrt{k}\right)^d (C_1/\ell)^{\frac{k-1}{2}} \|f\|_{L^2(\Omega)} + C_{\mathfrak{I}_\mathcal{T}} \alpha^{-1} \|Hf\|_{L^2(\Omega)}.$$

holds with C_1 from Lemma 3 and a constant C_4 that depends on C_H, β/α, and ρ but not on H, k, ℓ, f, 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 2, Lemma 4, and the triangle inequality,

$$\begin{aligned} |||u - u^{\mathrm{ms}}_{H,\ell k}|||^2 &\leq |||u - \tilde{u}^{\mathrm{ms}}_{H,\ell k}|||^2 = |||u - u^{\mathrm{ms}}_H + u^{\mathrm{ms}}_H - \tilde{u}^{\mathrm{ms}}_{H,\ell k}|||^2 \\ &\leq C^2_{\mathfrak{I}_\mathcal{T}} \alpha^{-2} \|Hf\|^2_{L^2(\Omega)} + C_3(\ell k)^d \sum_{x\in\mathcal{N}} u^{\mathrm{ms}}_H(x)^2 |||\phi_x - \phi_{x,\ell k}|||^2. \end{aligned} \tag{18}$$

The application of Lemma 3 yields

$$\sum_{x\in\mathcal{N}} u^{\mathrm{ms}}_H(x)^2 |||\phi_x - \phi_{x,\ell k}|||^2 \leq C_2^2 (C_1/\ell)^{k-1} \sum_{x\in\mathcal{N}} u^{\mathrm{ms}}_H(x)^2 |||\phi_x|||^2_{\omega_{x,\ell}}.$$

Observe that, $|||\sum_{y\in\mathcal{N}\cap\omega_{x,\ell}} u^{\mathrm{ms}}_H(y)\phi_y|||_{\omega_{x,L\ell}} = 0$ implies that $\sum_{y\in\mathcal{N}\cap\omega_{x,\ell}} u^{\mathrm{ms}}_H(y)\phi_y$ is constant on $\omega_{x,\ell}$. Hence, $\sum_{y\in\mathcal{N}\cap\omega_{x,L\ell}} u^{\mathrm{ms}}_H(y)\phi_y = \mathfrak{I}_\mathcal{T}\left(\sum_{y\in\mathcal{N}\cap\omega_{x,\ell}} u^{\mathrm{ms}}_H(y)\phi_y\right) = 0$ and there exists a constant C_4' such that

$$\sum_{y\in\mathcal{N}\cap\omega_{x,\ell}} u^{\mathrm{ms}}_H(y)^2 |||\phi_y|||^2_{\omega_{x,\ell}} \leq C_4' \left|\!\left|\!\left| \sum_{y\in\mathcal{N}\cap\omega_{x,\ell}} u^{\mathrm{ms}}_H(y)\phi_y \right|\!\right|\!\right|^2_{\omega_{x\ell}}.$$

The constant C_4' might depend on the shape regularity of $\mathcal{T}$ but a scaling argument shows that C_4' does not depend on the mesh size H and the parameter ℓ. Hence,

$$\begin{aligned} \sum_{x\in\mathcal{N}} u^{\mathrm{ms}}_H(x)^2 |||\phi_x|||^2_{\omega_{x,\ell}} &\leq \sum_{x\in\mathcal{N}} \sum_{y\in\mathcal{N}\cap\omega_{x,\ell}} u^{\mathrm{ms}}_H(y)^2 |||\phi_y|||^2_{\omega_{x,\ell}} \\ &\leq C_4' \sum_{x\in\mathcal{N}} \left|\!\left|\!\left| \sum_{y\in\mathcal{N}\cap\omega_{x,\ell}} u^{\mathrm{ms}}_H(y)\phi_y \right|\!\right|\!\right|^2_{\omega_{x,\ell}} \\ &\leq C_4''(\ell)^d |||u^{\mathrm{ms}}_H|||^2. \end{aligned}$$

The constant C_4'' depends only on C_4' and C_H. This yields

$$\begin{aligned} |||u^{\mathrm{ms}}_H - u^{\mathrm{ms}}_{H,\ell k}||| &\leq C_4(\ell^2 k)^{d/2} (C_1/\ell)^{(k-1)/2} |||u^{\mathrm{ms}}_H||| \\ &\leq C_4 \alpha^{-1} (\ell^2 k)^{d/2} (C_1/\ell)^{(k-1)/2} \|f\|_{L^2(\Omega)}, \end{aligned} \tag{19}$$

where C_4 only depends on C_2, C_3, and C_4''. The assertion follows readily be combining (18) and (19). ∎

4 Discretization of the Fine Scale Computations

We have not yet mentioned anything on how the local basis functions will be computed. There is a lot of freedom in choosing different finite elements and and different refinement strategies, see e.g. [14, 15]. We will here focus on a very simple and natural approach, also considered in [16]. 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 $\mathcal{T}$.

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

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

We let $u_h \in S_h \cap H_0^1(\Omega)$ be the reference solution solving,

$$a(u_h, v) = F(v) \quad \text{for all } v \in S_h \cap H_0^1(\Omega).$$

Locally on each patch we let $V_{h,x}(\omega_{x,k}) = S_h \cap H_0^1(\omega_{x,k})$ and furthermore,

$$V_{h,x}^{\mathrm{f}}(\omega_{x,k}) = \{v \in V_{h,x} \mid \Im_{\mathcal{T}} v = 0\}. \tag{20}$$

We now define the numerical approximation of $\phi_{x,k}^h$ in the following way: find $\phi_{x,k}^h \in V_{h,x}^{\mathrm{f}}(\omega_{x,k})$ such that,

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

We denote the discrete multiscale finite element space

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

The corresponding discrete multiscale approximation is given by: find $u_{H,k}^{\mathrm{ms},h} \in V_{H,k}^{\mathrm{ms},h}$ such that

$$a(u_{H,k}^{\mathrm{ms},h}, v) = F(v) \quad \text{for all } v \in V_{H,k}^{\mathrm{ms},h}. \tag{21}$$

Theorem 6 *Let $u \in V$ solve* (2) *and let $u_{H,\ell k}^{\mathrm{ms},h} \in V_{H,k}^{\mathrm{ms},h}$ solve* (21)*. Then*

$$|||u - u_{H,\ell k}^{\mathrm{ms},h}||| \leq C_4 \left(\ell \sqrt{k}\right)^d (C_1/\ell)^{\frac{k-1}{2}} \|f\|_{L^2(\Omega)} + C_{\Im_{\mathcal{T}}} \alpha^{-1} \|Hf\|_{L^2(\Omega)} + |||u - u_h|||.$$

Proof. We let $|||u-u_{H,\ell k}^{\mathrm{ms},h}||| \leq |||u-u_h||| + |||u_h-u_{H,\ell k}^{\mathrm{ms},h}|||$. The estimation of $|||u_h-u_{H,\ell k}^{\mathrm{ms},h}|||$ is almost verbatim the same as in the proof of Theorem 5 with $V := S_h \cap H_0^1(\Omega)$. The are only two minor issues.

First, the proof of Lemma 3 requires $\eta_j^2\phi_x \in S_h \cap V_f$ which requires a suitable choice of the cut-off function η_j. If $\tilde{\eta}_j$ is any cut-off function that satisfies (14), then η_j^2 can be defined as the S_h-piecewise rational function $\phi_x/\Im_{\mathcal{T}_h}\left(\tilde{\eta}_j^2\phi_x\right)$. Straight forward computations show that η_j satisfies (14). The fact that its definition depends on ϕ_x does not create additional difficulties.

Second, in Lemma 4 we need that $\zeta_x v_f$, with $v_f \in S_h \cap V_f$, stays in the the space $S_h \cap V_f$. We accomplish this by replacing $\zeta_x v_f$ with $(1-\Im_{\mathcal{T}})\Im_{\mathcal{T}_h}(\zeta_x v_f) \in S_h \cap V_f$. The argument then follows because $\Im_{\mathcal{T}}$ and $\Im_{\mathcal{T}_h}$ are both stable in H^1. ■

Remark 7 *The third part in the error bound in Theorem 6 can be bounded in terms of data, mesh parameter h, and polynomial degree p using standard a priori error estimates available for all properly analyzed finite element methods.*

Remark 8 *The local problems need to be solved in the spaces $V_{h,x}^{\mathrm{f}}(\omega_{x,k})$. This is a standard finite element space with the additional constraint that the trail 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, $dim(V_{H,k}^{ms,h}) = V_H$ 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.

5 Application to Multiscale Methods

Several multiscale methods for solving elliptic partial differential equations with heterogeneous coefficients have been developed during the last 15 years.

5.1 The Variational Multiscale Method

The variational multiscale method was first introduced in [12]. 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 then also decoupled into a coarse and a fine part. The method reads: find $\bar{u} \in S_H$ and $u' \in V^{\mathrm{f}}$ such that,

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

The fine scale solution is further decoupled over the coarse elements $T \in \mathcal{T}$ 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 method 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 symmetrical problem.

5.2 The Multiscale Finite Element Method

In [11] 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 basis functions fulfill the equations, find $\phi_{x,T} \in H_0^1(T)$

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

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 often referred to as over-sampling), see e.g. [11]. Note that since the coarse scale basis functions are modified (both trail and test space) the resulting method is symmetric.

5.3 The Proposed Method

The modified basis function construction given by equation (6-7) was first introduced in a variational multiscale framework in [13, 14]. It has also been used to modify basis functions in the spirit of the multiscale finite element method, see e.g. [17]. The exponential decay in terms of layers of coarse elements have been demonstrated numerically in several works, see e.g. [15, 17, 16]. From these papers it is clear that $l = 1$ or at most $l = 2$ in Theorem 5 is sufficient to get decay for all problems we have considered, i.e. $C_1 < 2$. The first paper where the exponential decay was analyzed is [16]. Here the theory of iterative solvers was used to show convergence to a discrete reference solution.

The proposed method has been extended to convection dominated problems and problems in mixed form [17] as well as time dependent problems [18]. A posteriori error bounds have been derived and adaptive algorithms designed where a local mesh and patch size are chosen automatically in order to reduce the error.

5.4 Application of the Presented Analysis

The convergence proof in this paper generalizes the results of [16] and gives a valid bound also as $h \to 0$ independent of the patch size. The proof does not rest

on regularity of the solution and gives a very explicit expression for the rate of convergence. The present analysis confirms the numerical results in [15, 17, 16] and gives together with [16] the method the solid theoretical foundation that has previously been missing. The analysis also justifies using the a posteriori error bounds for adaptivity [14, 17] since 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 e.g. [14, 17]. 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 our knowledge only very special cases, e.g. periodic coefficients, can then be analyzed.

References

[1] I. Babuška, G. Caloz, and J. E. Osborn. Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.*, 31(4):945–981, 1994.

[2] I. Babuška and R. Lipton. The penetration function and its application to microscale problems. *Multiscale Model. Simul.*, 9(1):373–406, 2011.

[3] I. Babuška and J. E. Osborn. Generalized finite element methods: their performance and their relation to mixed methods. *SIAM J. Numer. Anal.*, 20(3):510–536, 1983.

[4] I. Babuška and J. E. Osborn. Can a finite element method perform arbitrarily badly? *Math. Comp.*, 69(230):443–462, 2000.

[5] I. Babuška and J. M. Melenk. The partition of unity method. *International Journal of Numerical Methods in Engineering*, 40:727–758, 1996.

[6] L. Berlyand and H. Owhadi. Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast. *Arch. Ration. Mech. Anal.*, 198(2):677–721, 2010.

[7] C. Carstensen and S. A. Funken. Constants in Clément-interpolation error and residual based a posteriori error estimates in finite element methods. *East-West J. Numer. Math.*, 8(3):153–175, 2000.

[8] P. Ciarlet. *The finite element method for elliptic problems*. North-Holland, 1987.

[9] P. Clément. Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique*, 9(R-2):77–84, 1975.

[10] L. Grasedyck, I. Greff, and S. Sauter. The AL Basis for the Solution of Elliptic Problems in Heterogeneous Media. Technical report, Institut für Mathematik, Universität Zürich, Preprint 02-2011, 2011.

[11] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997.

[12] T. J. R. Hughes, G. R. Feijóo, L. Mazzei, and J.-B. Quincy. The variational multiscale method—a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166(1-2):3–24, 1998.

[13] M. G. Larson and A. Målqvist. Adaptive variational multiscale methods based on a posteriori error estimation: duality techniques for elliptic problems. In *Multiscale methods in science and engineering*, volume 44 of *Lect. Notes Comput. Sci. Eng.*, pages 181–193. Springer, Berlin, 2005.

[14] M. G. Larson and A. Målqvist. Adaptive variational multiscale methods based on a posteriori error estimation: energy norm estimates for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 196(21-24):2313–2324, 2007.

[15] M. G. Larson and A. Målqvist. A mixed adaptive variational multiscale method with applications in oil reservoir simulation. *Math. Models Methods Appl. Sci.*, 19(7):1017–1042, 2009.

[16] A. Målqvist. A priori error analysis of a multiscale method. *Preprint*.

[17] A. Målqvist. Multiscale methods for elliptic problems. *Multiscale Model. Simul.*, 9:1064–1086, 2011.

[18] J. M. Nordbotten. Adaptive variational multiscale methods for multiphase flow in porous media. *Multiscale Model. Simul.*, 7(3):1455–1473, 2008.

[19] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Wiley-Teubner series in advances in numerical mathematics. Wiley-Teubner, 1996.

[20] R. Verfürth. Error estimates for some quasi-interpolation operators. *M2AN Math. Model. Numer. Anal.*, 33(4):695–713, 1999.