A PRIORI ERROR ANALYSIS OF A MULTISCALE METHOD

AXEL MÅLQVIST

ABSTRACT. We derive an a priori error estimate and thereby prove convergence for the multiscale method presented by Larson and Målqvist in [13, 14]. The proof strongly relies on the local behavior of the elliptic differential operator on fine scales. The methodology can be extended and applied to other multiscale methods such as the multiscale finite element method [8]. We use iterative techniques to track down the decay rate of the fine scale basis functions for arbitrary positive bounded diffusion coefficient. The decay rate is the key result which leads to an a priori bound of the error in the multiscale approximation. We present five numerical test cases in order to illustrate the theoretical results of the paper.

1. INTRODUCTION

Various multiscale methods have been developed during the last two decades. Typical applications include e.g. porous media flow and mechanics of heterogeneous materials. A common feature for these applications is that they are very computationally expensive and often impossible to solve within an acceptable tolerance using standard one mesh methods. Multiscale methods, using local fine scale information in order to improve a coarse scale approximation, has turned out to be a very promising tool for tackling these difficulties.

1.1. **Previous work.** In [8] Hou and Wu present the multiscale finite element method, for solving elliptic partial differential equations with rapidly oscillating coefficients. The method is based on ideas from homogenization theory. Decoupled localized fine scale problems are solved on coarse elements, using a finer local subgrid, in order to modify the coarse scale basis functions. The coarse scale equations are then solved using the modified basis functions where the fine scale features are taken into account. To reduce the effect from the boundary conditions that are forced on the local problems, a method using larger subdomains, called over-sampling, has been introduced, see e.g. [6] and references therein. Error analysis for the multiscale finite element method is based on results from homogenization theory and is therefore restricted to very special cases, such as periodic coefficients, see e.g. [9].

The variational multiscale method (VMS) is an alternative approach which serves as a general framework for constructing multiscale methods, see [10, 11]. The idea is to decompose the solution into fine and coarse scale contributions, solve the fine scale equations

driven by the coarse scale residual, and finally eliminate the fine scale solution from the coarse scale equation. This procedure leads to a modified coarse scale equation where the modification accounts for the effect of fine scale behavior on the coarse scale. In several works various ways of analytical modeling have been investigated often based on bubbles or element Green's functions, see e.g. Arbogast [2] and Hughes [10]. Again the error analysis has concerned very special cases such as periodic coefficients.

The adaptive variational multiscale method (AVMS) was introduced by Larson and Målqvist in [13, 14]. The development was inspired by the solution of local problems on stars used in [18] to derive a posteriori error estimates and drive adaptivity. In AVMS the local problems are instead used to improve the solution. The fine scale equations are decoupled and solved numerically on subdomains that may be larger than the single mesh stars used in [18]. In [14] an a posteriori error estimate is presented, where the error is bounded in terms of coarse scale mesh size, fine scale mesh size, and size of the local subdomains. Based on this estimate an adaptive algorithm is constructed which automatically tunes the critical method parameters in order to meet the prescribed tolerance. This adaptive tuning of method parameters, which is the origin of the name AVMS, gives the method a very crucial advantage over competing methods. AVMS has been further developed in several directions: to convection diffusion problems [15], parabolic problems [20], and a mixed formulation of the Poisson equation [16]. Only a posteriori error estimates are presented in these papers.

A common advantage of the mentioned multiscale methods compared to more classical iterative methods for solving elliptic problems is that the multiscale methods often provide a modified coarse basis. Since the real applications are typically time dependent, often with small variations in the multiscale coefficient between different time steps, these basis functions can be computed once and then reused throughout the entire computation. Two phase flow simulation is a typical example of this when an elliptic equation is coupled to a hyperbolic equation and where the diffusion coefficient only varies at the front of the fluid as it travels through the domain. See [21] for more information about this model.

1.2. **New contributions.** The error analysis has so far, for the most part, been restricted to periodic coefficients while the multiscale methods have been applied to problems with non-periodic heterogeneities. The main exception is [14] but there only a posteriori error estimates are presented which unfortunately can not guarantee convergence of the method since the residuals present in the estimate are not easy to bound in terms of data.

In this paper we present an a priori error estimate for AVMS which also serve as a framework for error estimation of related methods such as the variational multiscale method and the multiscale finite element method. A key feature of AVMS is that we allow the support of the local subgrid solutions to grow in order to improve the convergence. This freedom is crucial when constructing a convergent multiscale method that can handle arbitrary diffusion coefficients. An error bound is presented in terms of the coaese scale mesh, the fine scale mesh size, and the size of the local subgrid domains. For simplicity we assume the same refinement and subgrid problem size for all local problems. It is indicated

how this can be generalized to also cover different resolutions and subgrid problem sizes in Remark 6.3. Numerical results for five different diffusion coefficients are provided.

1.3. **Outline.** In Section 2 we present some notations and preliminary results. In Section 3 we present the model problem and in Section 4 we present three different multiscale methods. In Section 5 we study approximation of the local basis functions of AVMS and in Section 6 we present the a priori error estimate, which is the main result of the paper. In Section 7 we present numerical examples and in Section 8 we draw conclusion and discuss future work.

2. Preliminaries

In this section we introduce notations frequently used in the paper. We also present some preliminary results.

2.1. **Functions spaces.** Let $\Omega \in \mathbf{R}^d$, where $d = 2, 3$, be a computational domain with polygonal boundary $\partial\Omega$. We let $\|\cdot\|_{L^2(\Omega)}$ be the $L^2(\Omega)$ norm of functions defined on Ω. We further let $\|\cdot\|_{L^\infty(\Omega)}$ denote the L^∞-norm on Ω, i.e. the essential supremum of the absolute value of the function over the domain Ω. We let $H^1(\Omega)$ denote the standard Sobolev space of functions in $L^2(\Omega)$ that has gradients in $L^2(\Omega)^d$. We let $D^2 v = (\sum_{i,j=1}^d |\frac{\partial^2 v}{\partial x_i \partial x_j}|^2)^{1/2}$ and $H^2(\Omega) = \{v \in H^1(\Omega) : \|D^2 v\|_{L^2(\Omega)} < \infty\}$. We further let $\mathcal{V} = H_0^1(\Omega)$ denotes the space of functions in $H^1(\Omega)$ with vanishing trace on the boundary $\partial\Omega$. See [1] for an extensive discussion on these function spaces. We are going to use a special notation for the energy scalar product and norm

$$(2.1) \qquad \langle v, w \rangle = \int_\Omega a \nabla v \cdot \nabla w \, dx \quad \text{and} \quad \|v\| = \langle v, v \rangle^{1/2}, \quad \forall v, w \in \mathcal{V}.$$

2.2. **Discrete function spaces.** Next we define meshes and discrete function spaces. We let $\mathcal{T}_0$ be a shape regular (with parameter γ bounding the diameter of the elements divided by the diameter of the largest inscribed ball, uniformly from above) coarse partition of Ω into a finite number of simplectic or quadrilateral elements τ, i.e., $\Omega = \bigcup_{\tau \in \mathcal{T}_0} \tau$. We generate $\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_j, \ldots$ by refining the initial partition $\mathcal{T}_0$ using either red-green or red local refinement strategies in $d = 2$ or $d = 3$ spatial dimensions. We let $\mathcal{N}_0$ be the set of interior nodes in the partition $\mathcal{T}_0$ and furthermore let $\mathcal{N}_j$ be the set of interior nodes in partition $\mathcal{T}_j$. We also let $\mathcal{M}_j = \mathcal{N}_j \backslash \mathcal{N}_0$, $j = 1, \ldots$, the mesh parameter $h_{j,\tau}$ be the diameter of $\tau \in \mathcal{T}_j$, and $h_j = \max_{\tau \in \mathcal{T}_j} h_{j,\tau}$, for all integers $j \geq 0$. Further, we let $\mathcal{V}_j$ be the (bi)linear continuous finite element functions defined on the mesh $\mathcal{T}_j$ equipped with zero boundary values. The successively refined finite element spaces, of continuous piecewise (bi)linear functions, will form the following nested sequence, $\mathcal{V}_0 \subset \mathcal{V}_1 \subset \cdots \subset \mathcal{V}_j \subset \cdots \subset \mathcal{V} = H_0^1(\Omega)$. We let $\pi_j : \mathcal{V} \cap C(\Omega) \to \mathcal{V}_0$ be the nodal interpolant and introduce the slice spaces $\mathcal{W}_j = \{v \in \mathcal{V}_j : \pi_0 v = 0\}$ of functions in $\mathcal{V}_j$ but with no component in $\mathcal{V}_0$, i.e. $\mathcal{V}_j = \mathcal{V}_0 \oplus \mathcal{W}_j$. We let $C(\Omega)$ denote the space of continuous functions on Ω. We note that the degrees of freedom in $\mathcal{W}_J$ are located at the nodes in $\mathcal{M}_J$.

2.3. **Matrices and vectors.** Let $v \in \mathbf{R}^n$ be a given vector. We let $|v| = \left(\sum_{i=1}^n v_i^2\right)^{1/2}$ be the Euclidian norm of v. Given a matrix $M \in \mathbf{R}^{n \times n}$ we let $\lambda_{\min}(M) = \min_{v \neq 0} \frac{v^T M v}{v^T v}$ and $\lambda_{\max}(M) = \max_{v \neq 0} \frac{v^T M v}{v^T v}$ where v^T denotes the transpose of v, and

$$\kappa(M) = \frac{\lambda_{\max}(M)}{\lambda_{\min}(M)},$$

be the condition number of the matrix M. For a symmetric positive definite matrix A we also introduce the $|\cdot|_A$-norm as, $|v|_A = (v^T A v)^{1/2}$ for all $v \in \mathbf{R}^n$. We note that $\lambda_{\min}(A)|v|^2 \leq |v|_A^2 \leq \lambda_{\max}(A)|v|^2$.

In this paper we also use the conjugated gradient method in the analysis. We present the standard convergence result below.

Lemma 2.1. *Given a symmetric positive definite matrix $A \in \mathbf{R}^{n \times n}$ and a vector $b \in \mathbf{R}^n$ the conjugated gradient iterates x^k fulfill,*

$$(2.2) \qquad |x - x^k|_A \leq 2\rho^k |x - x^0|_A, \quad where \ \rho = \frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1},$$

for a given initial guess $x^0 \in \mathbf{R}^n$.

Proof. See e.g. Theorem 10.2.6 in [7]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Finally we present a bound for $\kappa(\hat{A})$ for a certain matrix $\hat{A}$ that will be used in this paper. Let $\{\chi_i\}_{i \in \mathcal{M}_J}$ be a hierarchical basis for the space $\mathcal{W}_J$ and let $\hat{A}$ be the matrix with entries $\hat{A}_{i,j} = \langle \chi_i, \chi_j \rangle$, for all $i, j \in \mathcal{M}_J$. Then,

$$(2.3) \qquad \kappa(\hat{A}) \leq \begin{cases} C_a \left(\log \dfrac{h_0}{h_J}\right)^2, & \text{for } d = 2, \\[2ex] C_a \left(\dfrac{h_0}{h_J}\right)^2, & \text{for } d = 3, \end{cases}$$

where C_a depends on a. See e.g. Theorem 5.8 in [17].

3. Model problem

We consider the Poisson equation with homogeneous Dirichlet boundary conditions,

$$(3.1) \qquad \begin{cases} -\nabla \cdot a \nabla u = f, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases}$$

where the diffusion coefficient a has multiscale features and fulfills $0 < a_0 \leq a \in L^\infty(\Omega)$ for some $a_0 \in \mathbf{R}$. The right hand side $f \in L^2(\Omega)$. We introduce a linear functional $l : \mathcal{V} \to \mathbf{R}$,

$$(3.2) \qquad l(v) = \int_\Omega f v \, dx, \quad \text{for all } v \in \mathcal{V}.$$

The weak form of equation (3.1) reads: find $u \in \mathcal{V}$ such that,

$$(3.3) \qquad \langle u, v \rangle = l(v), \quad \text{for all } v \in \mathcal{V}.$$

In this paper we aim at computing an approximation of the finite element solution on the reference mesh $\mathcal{T}_J$. We therefore refer to $u_J \in \mathcal{V}_J$ which solves,

$$(3.4) \qquad \langle u_J, v \rangle = l(v), \quad \text{for all } v \in \mathcal{V}_J,$$

as the reference solution. The orthogonality between the error $u - u_J$ and the space $\mathcal{V}_J$ in the $\langle \cdot, \cdot \rangle$ scalar product reduces the problem of estimating the error in the finite element solution to approximation theory for the finite element. By subtracting equation (3.3) from (3.4) and using the Cauchy-Schwarz inequality we get the following identity,

$$(3.5) \qquad \|u - u_J\| = \min_{v \in \mathcal{V}_J} \|u - v\|.$$

Remark 3.1 If $u \in H^2(\Omega) \cap H_0^1(\Omega)$ we have the following a priori error bound, see e.g. [4],

$$(3.6) \qquad \|u - u_J\| \leq \min_{v \in \mathcal{V}_J} \|u - v\| \leq C_{a,\Omega} h_J \|D^2 u\|_{L^2(\Omega)}.$$

However, when the coefficient a varies rapidly $C_{a,\Omega}$ will depend on the inverse of the scale on which a varies. This means that we will have to resolve the fine scale features in a with the mesh in order to get a reliable solution, see e.g. equation (4.6) in [9].

4. Multiscale methods

We aim at computing an approximation in a coarse subspace of $\mathcal{V}_J$ (with $n = \dim(\mathcal{N}_0)$ degrees of freedom) to the reference solution u_J defined in equation (3.4). We decompose the space $\mathcal{V}_J$ into a coarse and a fine part using the nodal interpolant π_0, we let $\mathcal{V}_J = \mathcal{V}_0 \oplus \mathcal{W}_J$, where $\mathcal{W}_J = \{v \in \mathcal{V}_J : \pi_0 v = 0\}$. In order to separate the scales we introduce an operator $T_J : \mathcal{V}_0 \to \mathcal{W}_J$ in the following way, for any $v_0 \in \mathcal{V}_0$, let

$$(4.1) \qquad \langle v_0 + T_J v_0, w \rangle = 0, \quad \text{for all } w \in \mathcal{W}_J.$$

This means that we have access to a basis $v_0 + T_J v_0$ of functions that are orthogonal to the space $\mathcal{W}_J$ in the $\langle \cdot, \cdot \rangle$ scalar product. We now split u_J in two parts, $u_J = (u_0 + T_J u_0) + u_{l,J}$, where $u_0 = \pi_0 u_J$, $T_J u_0$ is defined by equation (4.1), and $u_{l,J} = (1 - \pi_0 - T_J \pi_0) u_J$. We note that $u_{l,J} \in \mathcal{W}_J$ solves,

$$(4.2) \qquad \langle u_{l,J}, w \rangle = l(w), \quad \text{for all } w \in \mathcal{W}_J.$$

By plugging in $u_0 + T_J u_0$ and $u_{l,J}$ in equation (3.4) we get the following equation, find $u_0 \in \mathcal{V}_0$ such that,

$$(4.3) \qquad \langle u_0 + T_J u_0, v \rangle = l(v) - \langle u_{l,J}, v \rangle, \quad \text{for all } v \in \mathcal{V}_J.$$

Starting from equation (4.3) we can choose v and approximate T_J, and $u_{l,J}$ in different ways to derive different multiscale methods.

4.1. **The variational multiscale method.** Here we let $v = v_0 \in \mathcal{V}_0$ and keep $u_{l,J}$ in the right hand side of equation (4.3),

$$(4.4) \qquad \langle u_0 + T_J^{\mathrm{vms}} u_0, v_0 \rangle = l(v_0) - \langle u_{l,J}^{\mathrm{vms}}, v_0 \rangle, \quad \text{for all } v_0 \in \mathcal{V}_0,$$

$$(4.5) \qquad \langle v_0 + T_J^{\mathrm{vms}} v_0, v \rangle \approx 0, \quad \text{for all } v \in \mathcal{W}_J,$$

$$(4.6) \qquad \langle u_{l,J}^{\mathrm{vms}}, v \rangle \approx l(v), \quad \text{for all } v \in \mathcal{W}_J.$$

The approximation sign $\approx$ indicates that this is computed approximately. Typically the fine scale equations (4.5-4.6) are decoupled over each coarse element and approximations to T_J and $u_{l,J}$ are derived using analytical techniques, see e.g. [10]. Note that the bilinear form $\langle u_0 + T_J^{\mathrm{vms}} u_0, v_0 \rangle$ is non-symmetric unlike the original problem. The approximate solution is given by $u_J^{\mathrm{vms}} = u_0 + T_J^{\mathrm{vms}} u_0 + u_{l,J}^{\mathrm{vms}} \approx u_J$.

4.2. **The multiscale finite element method.** In the multiscale finite element method $v = v_0 + T_J v_0$, with $v_0 \in \mathcal{V}_0$, and the identity $\langle u_{l,J}, v_0 + T_J v_0 \rangle = 0$ is used to drop the second part of the right hand side in equation (4.3), see e.g. [8],

$$(4.7) \qquad \langle u_0 + T_J^{\mathrm{mfem}} u_0, v_0 + T_J^{\mathrm{mfem}} v_0 \rangle = l(v_0 + T_J^{\mathrm{mfem}} v_0), \quad \text{for all } v_0 \in \mathcal{V}_0,$$

$$(4.8) \qquad \langle v_0 + T_J^{\mathrm{mfem}} v_0, v \rangle \approx 0, \quad \text{for all } v \in \mathcal{W}_J.$$

The fine scale equations (4.8) are decoupled and solved numerically either on one coarse element or on a larger domain surrounding a coarse element (over-sampling) and then restricted back to the coarse element. We note that the bilinear form is symmetric just like the original problem. In this approach the approximation is $u_J^{\mathrm{mfem}} = u_0 + T_J^{\mathrm{mfem}} u_0 \approx u_J$ and an approximation to $u_{l,J}$ is never computed.

4.3. **The proposed method.** In [13, 14] AVMS was first presented. This approach can be viewed as a combination of these two prior formulations combined with an adaptive algorithm that automatically decides how accurate the local subgrid problems need to be solved to meet an overall tolerance of the global error. This adaptive algorithm is based on an a posteriori error estimate. For practical reasons we refer to the approximation of T_J and $u_{l,J}$ as T_J^k and $u_{l,J}^k$ rather than the more natural T_J^{avms} and $u_{l,J}^{\mathrm{avms}}$. We have,

$$(4.9) \qquad \langle u_0 + T_J^k u_0, v_0 + T_J^k v_0 \rangle = l(v_0 + T_J^k v_0) - \langle u_{l,J}^k, v_0 + T_J^k v_0 \rangle, \text{ for all } v_0 \in \mathcal{V}_0,$$

$$(4.10) \qquad \langle v_0 + T_J^k v_0, v \rangle \approx 0, \quad \text{for all } v \in \mathcal{W}_J,$$

$$(4.11) \qquad \langle u_{l,J}^k, v \rangle \approx l(v), \quad \text{for all } v \in \mathcal{W}_J.$$

Here we have picked $v = v_0 + T_J^k v_0$, where $v_+ \in \mathcal{V}_0$, to get a symmetric formulation and we have included the right hand side term $\langle u_{l,J}^k, v_0 + T_J^k v_0 \rangle$. Here the approximation is, $u_J^k = u_0 + T_J^k u_0 + u_{l,J}^k \approx u_J$. In the next section we will explain exactly how T_J^k and $u_{l,J}^k$ are computed.

Remark 4.1 We note that $\langle u_{l,J}, v_0 + T_J v_0 \rangle = 0$ which indicates that this term might be skipped also in the approximate formulations. However, the approximations of T_J do not have full orthogonality and it appears from the error analysis in Section 6 that this term
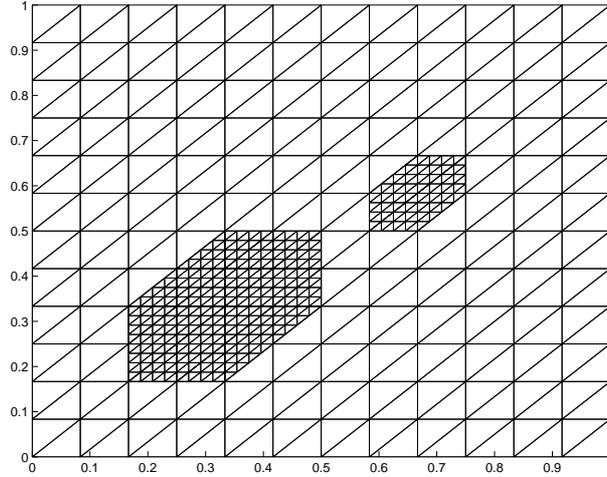
FIGURE 1. A 2-ring (left) and a 1-ring (right) of coarse elements on a structured triangular grid. Here $\Omega = [0,1] \times [0,1]$, $J = 2$, and $h_0 = \sqrt{2}/12$.

is needed to prove convergence.

Remark 4.2 We note that the term $u_{l,J}$ is not included in the approximation for the multiscale finite element method as it is described in [8]. This only makes sense if the fine scale part of f is small since

$$(4.12) \qquad \|u_{l,J}\| = \sup_{v \in \mathcal{W}_J : \|v\|=1} \langle u_{l,J}, v \rangle = \sup_{v \in \mathcal{W}_J : \|v\|=1} \int_\Omega f v \, dx = \sup_{v \in \mathcal{V} : \|v\|=1} \int_\Omega f Q_{0,J} v \, dx$$

$$(4.13) \qquad = \sup_{v \in \mathcal{V} : \|v\|=1} \int_\Omega Q_{0,J} f v \, dx := \|Q_{0,J} f\|_{\mathcal{V}^*},$$

where $Q_{0,J} : L^2(\Omega) \to \mathcal{W}_J$ is the L^2-projection onto $\mathcal{W}_J$ and $\|\cdot\|_{\mathcal{V}^*}$ defined in equation (4.13) is a norm in the dual space $\mathcal{V}^*$. Since $u_J - u_0 - T_J u_0 = u_{l,J}$ we can not expect a small error even with the exact T_J unless $Q_{0,J} f$ has a small $\mathcal{V}^*$ norm.

5. APPROXIMATION OF FINE SCALE SOLUTIONS

Equations (4.1) and (4.2) need to be solved approximately. The first step is to localize the computations by using the partition of unity made out of coarse basis functions,

$$(5.1) \qquad \langle \phi_i + T_J \phi_i, w \rangle = 0, \quad \text{for all } w \in \mathcal{W}_J,$$

and all $i \in \mathcal{N}_0$, where $\mathcal{V}_0 = \text{span}(\{\phi_i\}_{i \in \mathcal{N}_0})$. Equation (5.1) is then solved approximately on a k-ring ω_i^k (see Figure 1) of coarse elements surrounding node x_i (for which $\phi_i(x_i) = 1$), using homogeneous Dirichlet boundary conditions. We let,

$$(5.2) \qquad \mathcal{W}_J(\omega_i^k) = \{v \in \mathcal{W}_J : \text{supp}(v) \subset \omega_i^k\},$$

and define approximations $T_J^k \phi_i \in \mathcal{W}_J(\omega_i^k) \subset \mathcal{W}_J$ and $u_{l,J}^k = \sum_{i \in \mathcal{N}_0} u_{l,J,i}^k$, where $u_{l,J,i}^k \in \mathcal{W}_J(\omega_i^k) \subset \mathcal{W}_J$, that fulfills,

$$(5.3) \qquad \langle \phi_i + T_J^k \phi_i, w \rangle = 0, \quad \text{for all } w \in \mathcal{W}_J(\omega_i^k),$$

$$(5.4) \qquad \langle u_{l,J,i}^k, v \rangle = l(\phi_i v), \quad \text{for all } v \in \mathcal{W}_J(\omega_i^k).$$

Note that the problems are totally decoupled. For an arbitrary $v_0 = \sum_{i \in \mathcal{N}_0} v_0^i \phi_i \in \mathcal{V}_0$ we define, $T_J^k v_0 = \sum_{i \in \mathcal{N}_0} v_0^i T_J^k \phi_i$.

5.1. **The proposed method formulated using projections.** We now have two modified sets of basis functions, defined by equations (5.1) and (5.3), with the same dimension as the coarse space $\mathcal{V}_0$. We introduce the following notations for these spaces,

$$(5.5) \qquad \mathcal{V}_{0,J} = \text{span}(\{\phi_i + T_J \phi_i\}_{i \in \mathcal{N}_0}),$$

$$(5.6) \qquad \mathcal{V}_{0,J}^k = \text{span}(\{\phi_i + T_J^k \phi_i\}_{i \in \mathcal{N}_0}),$$

and formulate the reference solution and the multiscale approximation as projections, $P_J : \mathcal{V} \to \mathcal{V}_{0,J}$ and $P_J^k : \mathcal{V} \to \mathcal{V}_{0,J}^k$ in the following way. Note that $\langle u_0^k + T_J^k u_0^k, v \rangle = l(v) - \langle u_{l,J}^k, v \rangle = \langle u_J - u_{l,J}^k, v \rangle$, where $u_0^k + T_J^k u_0^k = u_J^k - u_{l,J}^k \in \mathcal{V}_{0,J}^k$ so $u_0^k + T_J^k u_0^k = P_J^k(u_J^k - u_{l,J}^k)$, i.e.,

$$(5.7) \qquad \langle P_J u_J, v \rangle = \langle u_J, v \rangle, \quad \text{for all } v \in \mathcal{V}_{0,J},$$

$$(5.8) \qquad \langle P_J^k(u_J - u_{l,J}^k), v \rangle = \langle u_J - u_{l,J}^k, v \rangle, \quad \text{for all } v \in \mathcal{V}_{0,J}^k.$$

We note that the orthogonality between the space $\mathcal{V}_{0,J}$ and the function $u_{l,J}$ makes the formulation of the reference projection P_J different from its approximation. We also note that, $u_J = P_J u_J + u_{l,J}$ and that the computable approximation of u_J is given by,

$$(5.9) \qquad u_J^k = P_J^k(u_J - u_{l,J}^k) + u_{l,J}^k.$$

The function $P_J^k(u_J - u_{l,J}^k)$ is clearly computable since $u_{l,J}^k$ and $\{\phi_i + T_J^k \phi_i\}_{i=1}^n$ are computable and, $\langle P_J^k(u_J - u_{l,J}^k), v \rangle = l(v) - \langle u_{l,J}^k, v \rangle$ for all $v \in \mathcal{V}_{0,J}^k$. Existence and uniqueness follows from the Lax-Milgram Lemma since $\mathcal{V}_{0,J}^k \subset \mathcal{V}_J$ and the bilinear form clearly is coercive and bounded by assumption.

6. A PRIORI ERROR ANALYSIS

In this section we prove an a priori error bound for the proposed method, i.e. we bound $u_J - u_J^k$ in terms of h_0, J, k, and computable constants. This bound can then easily be combined with a classical a priori bound for the error $u - u_J$, see equation (3.6), in order to get a final bound of the true error $u - u_J^k$. Before we present the proof in two technical Lemma's and a main Theorem followed by a Corollary we present a short summary of the main argument:

   (1) In Lemma 6.1 we control the decay of $T_J \phi_i$ away form the support of ϕ_i. The main idea is the following. The right hand side of the local problems has support in a coarse 1-ring. If we use an iterative methods like the conjugated gradient algorithm and furthermore use an hierarchical split of $\mathcal{W}_J$ when we compute approximations

to $T_J\phi_i$ we note that the support of the approximation after $2k$ iterations will be a k-ring of coarse elements and the error in the approximation will be bounded by $2\rho^{2k}$, where ρ is small because of the hierarchical basis. Note that the the conjugated gradient method is used strictly as an analytical tool and is *not* used in the numerical method. The procedure described gives a bound of $T_J\phi_i$ outside the k-ring and thereby control over the decay.

(2) Furthermore in Lemma 6.1 we bound the error in $\|T_J\phi_i - T_J^k\phi_i\|$ by bounding this quantity in terms of the values of $T_J\phi_i$ outside the k-ring. We also bound the error in $\|u_{l,J,i} - u_{l,J,i}^k\|$ using the same argument as above.

(3) In Lemma 6.2 and Theorem 6.1 we bound the error $\|u_J - u_J^k\|$ by viewing the two solutions as projections onto different subspaces, with perturbed basis functions. The error in the perturbed basis is given by item 2.

(4) Finally in Corollary 6.1 we simply combine equation (3.6) and Theorem 6.1 to get a bound of $\|u - u_J^k\|$.

Lemma 6.1. *Let $\{\chi_k\}_{k\in\mathcal{M}_J}$ be a hierarchical basis for $\mathcal{W}_J$ and $\hat{A}$ be the matrix with entries $\hat{A}_{k,j} = \langle\chi_k,\chi_j\rangle$ for all $j,k\in\mathcal{M}_J$. It holds*

$$\|T_J\phi_i - T_J^m\phi_i\| \leq 4\kappa(\hat{A})^{1/2}\rho^{2m}\|T_J\phi_i\|, \tag{6.1}$$

$$\|u_{l,J,i} - u_{l,J,i}^m\| \leq \frac{4C_{PF}\kappa(\hat{A})^{1/2}\rho^{2m}}{a_0^{1/2}}\|f\phi_i\|_{L^2(\Omega)}, \tag{6.2}$$

where $\rho = \frac{\sqrt{\kappa(\hat{A})}-1}{\sqrt{\kappa(\hat{A})}+1}$, $m = 1,2,\ldots$, and C_{PF} is a Poincare-Friedrich constant further discussed in Remark 6.1.

Proof. Let $\mathcal{M}_J(\omega) = \{k \in \mathcal{M}_J : \text{supp}(\chi_k) \subset \omega\}$ for an arbitrary $\omega \in \Omega$ and let $\mathcal{M}_J(\omega_1 \setminus \omega_2) = \mathcal{M}_J(\omega_1) \setminus \mathcal{M}_J(\omega_2)$ for arbitrary sets $\omega_2 \subset \omega_1 \subset \Omega$. We let $T_J\phi_i = \sum_{k\in\mathcal{M}_J}\alpha_k\chi_k$. We further introduce vectors (of length $\dim(\mathcal{M}_J)$) α and α_ω corresponding to coefficients $\{\alpha_k\}_{\mathcal{M}_J}$ and $\{\alpha_k\}_{\mathcal{M}_J(\omega)}$ for an arbitrary set ω, i.e. we add zeros to the entries in the vector that are outside the set $\mathcal{M}_J(\omega)$.

We use the conjugated gradient method to approximate α, let $\hat{\alpha}^l$ be the l:th iterate with initial guess $\hat{\alpha}^0 = 0$, matrix $\hat{A}$ and right hand side $b_k = -\langle\phi_i,\chi_k\rangle$. By convergence analysis from Lemma 2.1 we have after m iterations,

$$|\alpha - \hat{\alpha}^m|_{\hat{A}} \leq 2\left(\frac{\sqrt{\kappa(\hat{A})}-1}{\sqrt{\kappa(\hat{A})}+1}\right)^m |\alpha|_{\hat{A}} := 2\rho^m|\alpha|_{\hat{A}}. \tag{6.3}$$

Since the right hand side b_k has support on a coarse 1-ring surrounding node i and the hierarchical basis $\{\chi_k\}$ only can spread information within ω_i^m in $2m$ iterations, we conclude $\sum_{k\in\mathcal{M}_J(\Omega\setminus\omega_i^m)}\hat{\alpha}_k^{2m}\chi_k = 0$. We have,

$$|\alpha_{\Omega\setminus\omega^m}|^2 = \sum_{k\in\mathcal{M}_J(\Omega\setminus\omega_i^m)}|\alpha_k|^2 = \sum_{k\in\mathcal{M}_J(\Omega\setminus\omega_i^m)}|\alpha_k - \hat{\alpha}_k^{2m}|^2 \leq |\alpha - \hat{\alpha}^{2m}|^2. \tag{6.4}$$

In the $|\cdot|_{\hat{A}}$ norm we get, $|\alpha_{\Omega\setminus\omega^m}|^2_{\hat{A}} \leq \kappa(\hat{A})|\alpha - \hat{\alpha}^{2m}|^2_{\hat{A}} \leq 4\kappa(\hat{A})\rho^{4m}|\alpha|^2_{\hat{A}}$.

Furthermore we let $T_J^m\phi_i = \sum_{k\in\mathcal{M}_J}\alpha_k^m\chi_k$, with corresponding $\dim(\mathcal{M}_J)$ vector α^m (with zeros in $\mathcal{M}_J(\Omega) \setminus \mathcal{M}_J(\omega_i^m)$). We note that,

$$(6.5) \qquad |\alpha_{\omega^m} - \alpha^m|^2_{\hat{A}} = -\alpha^T_{\Omega\setminus\omega^m}\hat{A}(\alpha_{\omega^m} - \alpha^m) \leq |\alpha_{\Omega\setminus\omega^m}|_{\hat{A}}|\alpha_{\omega^m} - \alpha^m|_{\hat{A}},$$

since the function corresponding to $\alpha_{\omega^m} - \alpha^m$ is in $\mathcal{W}_J(\omega_i^m)$ and the definition of the multiscale basis functions (5.1,5.3) gives $w^T\hat{A}(\alpha-\alpha^m) = 0$ for all vectors with corresponding function in the same space $\sum_{i\in\mathcal{M}_J}w_i\chi_i \in \mathcal{W}_J(\omega_i^m)$, (where $w = [w_1, w_2, \ldots]$).

We get $|\alpha_\omega - \alpha^m|^2_{\hat{A}} \leq 4\kappa(\hat{A})\rho^{4m}|\alpha|^2_{\hat{A}}$ and therefore,

$$(6.6) \quad |\alpha - \alpha^m|_{\hat{A}} = |\alpha_{\omega^m} - \alpha^m + \alpha_{\Omega\setminus\omega^m}|_{\hat{A}} \leq |\alpha_{\omega^m} - \alpha^m|_{\hat{A}} + |\alpha_{\Omega\setminus\omega^m}|_{\hat{A}} \leq 4\kappa(\hat{A})^{1/2}\rho^{2m}|\alpha|_{\hat{A}}$$

or,

$$(6.7) \qquad \|T_J\phi_i - T_J^m\phi_i\| = |\alpha - \alpha^m|_{\hat{A}} \leq 4\kappa(\hat{A})^{1/2}\rho^{2m}\|T_J\phi_i\|.$$

The first result of the Lemma is thereby proven. For the second result we note that the exact same argument also applies for $u_{l,J,i} = \sum_{k\in\mathcal{M}_J}\beta_k\chi_k$ with $b_k = \int_\Omega f\phi_i\chi_k\,dx$. We conclude, using similar notation as above, that

$$(6.8) \qquad \|u_{l,J,i} - u_{l,J,i}^m\| \leq |\beta_{\omega^m} - \beta^m|_{\hat{A}} + |\beta_{\Omega\setminus\omega^m}|_{\hat{A}} \leq 4\kappa(\hat{A})^{1/2}\rho^{2m}\|u_{l,J,i}\|.$$

Finally, we have $\|u_{l,J,i}\|^2 = \int_\Omega f\phi_i u_{l,J,i}\,dx \leq \frac{C_{\text{PF}}}{\sqrt{a_0}}\|f\phi_i\|_{L^2(\Omega)}\|u_{l,i,J}\|$, using the Poincare-Friedrich inequality and that a is bounded from below. The second part of the Lemma now follows immediately. $\qquad\square$

Lemma 6.2. *Given the references solution u_J defined in equation (3.4), the projections P_J and P_J^k defined in equations (5.7-5.8), and the hierarchical matrix $\hat{A}$ defined in the statement of Lemma 6.1, it holds,*

$$(6.9) \qquad \|P_J u_J - P_J^k P_J u_J\| \leq 4C_{inv,a}(\max_{\tau\in\mathcal{T}_0}h_{0,\tau}^{-1})\kappa(\hat{A})^{1/2}\|u_J\|_{L^\infty(\Omega)}\rho^{2k},$$

where $k = 1, 2, \ldots$ and $\rho = \frac{\sqrt{\kappa(\hat{A})}-1}{\sqrt{\kappa(\hat{A})}+1}$.

Proof. Let $P_J u_J = \sum_{i\in\mathcal{N}_0}v_i(\phi_i + T_J\phi_i) \in \mathcal{V}_{0,J}$ and $w = \sum_{i\in\mathcal{N}_0}v_i(T_J\phi_i - T_J^k\phi_i) \in \mathcal{V}_{0,J}^k$. Since P_J^k is a projection (in the $\langle\cdot,\cdot\rangle$ scalar product) we have that $\|(1 - P_J^k)P_J u\| \leq \|w\|$, and furthermore,

$$(6.10) \qquad \|w\|^2 \leq \sum_{i\in\mathcal{N}_0}|v_i|^2\|T_J\phi_i - T_J^k\phi_i\|^2$$

$$(6.11) \qquad \leq 16\max_{i\in\mathcal{N}_0}|v_i|^2\kappa(\hat{A})\rho^{4k}\sum_{i\in\mathcal{N}_0}\|T_J\phi_i\|^2$$

$$(6.12) \qquad \leq 16\max_{i\in\mathcal{N}_0}|v_i|^2\kappa(\hat{A})\rho^{4k}\sum_{i\in\mathcal{N}_0}\|\phi_i\|^2$$

$$(6.13) \qquad \leq 16\max_{i\in\mathcal{N}_0}|v_i|^2\kappa(\hat{A})C_{\text{inv},a}^2(\max_{\tau\in\mathcal{T}_0}h_{0,\tau}^{-2})\rho^{4k},$$

where we have used that $\sum_{i\in\mathcal{N}_0}\|\phi_i\|^2 \leq C_{\mathrm{inv},a}^2 \max_{\tau\in\mathcal{T}_0} h_{0,\tau}^{-2}$ for a constant $C_{\mathrm{inv},a}$ further discussed in Remark 6.1. We note that $\max_{i\in\mathcal{N}_0}|v_i| = \|\pi_0 u_J\|_{L^\infty(\Omega)} \leq \|u_J\|_{L^\infty(\Omega)}$ since $\{v_i\}_{i\in\mathcal{N}_0}$ are nodal values of u_J (in coarse nodes). We conclude,

$$(6.14) \qquad \|P_J u - P_J^k P_J u\| \leq \|w\| \leq 4C_{\mathrm{inv},a}(\max_{\tau\in\mathcal{T}_0} h_{0,\tau}^{-1})\kappa(\hat{A})^{1/2}\|u_J\|_{L^\infty(\Omega)}\rho^{2k}.$$

$\square$

We are ready to present the main theorem.

Theorem 6.1. *Let u_J be the reference solution defined in equation (3.4) and let $u_J^k = P_J^k(u_J - u_{l,J}^k) - u_{l,J}^k$ defined in equation (5.9). Then,*

$$(6.15) \qquad \|u_J - u_J^k\| \leq \left(4C_{inv,a}(\max_{\tau\in\mathcal{T}_0} h_{0,\tau}^{-1})\|u_J\|_{L^\infty(\Omega)} + \frac{2\sqrt{32d}C_{PF}}{\sqrt{a_0}}\|f\|_{L^2(\Omega)}\right)\kappa(\hat{A})^{1/2}\rho^{2k},$$

where $\rho = \frac{\sqrt{\kappa(\hat{A})}-1}{\sqrt{\kappa(\hat{A})}+1}$.

Proof. We have $u_J^k = P_J^k(u_J - u_{l,J}^k) - u_{l,J}^k$ and,

$$(6.16) \quad \|u_J - P_J^k(u_J - u_{l,J}^k) - u_{l,J}^k\| \leq \|P_J u_J - P_J^k(u_J - u_{l,J}^k)\| + \|u_{l,J} - u_{l,J}^k\| \leq e_\mathrm{I} + e_\mathrm{II}.$$

For e_II we apply Lemma 6.1 in the following calculation,

$$(6.17) \qquad e_\mathrm{II}^2 \leq \sum_{i\in\mathcal{N}_0} \|u_{l,J,i} - u_{l,J,i}^k\|^2$$

$$(6.18) \qquad \leq \sum_{i\in\mathcal{N}_0} \frac{16C_\mathrm{PF}^2\kappa(\hat{A})\rho^{4k}}{a_0}\|f\phi_i\|_{L^2(\Omega)}^2$$

$$(6.19) \qquad \leq \frac{16C_\mathrm{PF}^2\kappa(\hat{A})\rho^{4k}}{a_0} \sum_{i\in\mathcal{N}_0} \|f\|_{L^2(\mathrm{supp}(\phi_i))}^2$$

$$(6.20) \qquad \leq \frac{16\cdot 2dC_\mathrm{PF}^2\kappa(\hat{A})\rho^{4k}}{a_0}\|f\|_{L^2(\Omega)}^2,$$

where $2d$ can be replaced with $(d+1)$ for simplectic elements ($2d$ is sharp for quadrilateral elements), since each element in $\mathcal{T}_0$ gets counted $(d+1)$ and $2d$ times respectively.

We consider e_I and use that $I = (I - P_J^k) + P_J^k$ and that $\|P_J^k v\| \leq \|v\|$, for all $v \in \mathcal{V}_J$, to get:

$$(6.21) \quad \|P_J u - P_J^k(u_J - u_{l,J}^k)\| \leq \|(I - P_J^k)P_J u_J\| + \|P_J^k P_J u_J - P_J^k(u_J - u_{l,J}^k)\|$$

$$(6.22) \qquad \leq \|(I - P_J^k)P_J u_J\| + \|P_J u_J - u_J + u_{l,J}\| + \|u_{l,J} - u_{l,J}^k\|$$

$$(6.23) \qquad \leq \|(I - P_J^k)P_J u_J\| + 0 + e_\mathrm{II},$$

where we use the same definition of e_II as above. Again we can apply equation (6.20) and for the first term $\|(I - P_J^k)P_J u_J\|$ we apply Lemma 6.2. The theorem follows immediately. $\square$

Corollary 6.1. *Let u be the reference solution defined in equation (3.3), such that $u \in H_0^1(\Omega) \cap H^2(\Omega)$, and let $u_J^k = P_J^k(u_J - u_{l,J}^k) - u_{l,J}^k$ defined in equation (5.9). Then,*

$$(6.24) \quad \|u - u_J^k\| \leq C_{a,\Omega} h_J \|D^2 u\|_{L^2(\Omega)}$$

$$(6.25) \qquad\qquad + \left(4 C_{inv,a} (\max_{\tau \in \mathcal{T}_0} h_{0,\tau}^{-1}) \|u_J\|_{L^\infty(\Omega)} + \frac{2\sqrt{32d} C_{PF}}{\sqrt{a_0}} \|f\|_{L^2(\Omega)} \right) \kappa(\hat{A})^{1/2} \rho^{2k},$$

where $k = 1, 2, \ldots$ and $\rho = \frac{\sqrt{\kappa(\hat{A})} - 1}{\sqrt{\kappa(\hat{A})} + 1}$ and $C_{a,\Omega}$ is a constant depending on a and Ω.

Proof. Follows immediately by combining equation (3.6) and Theorem 6.1. $\qquad\square$

Both in Theorem 6.1 and Corollary 6.1 we note that the error increases as J increases or h_0 decreases with constant J. This means that more layers of coarse elements in the k-rings, i.e. a higher k, will be needed as the mesh gets refined. This should be compared with iterative methods where more iterative are typically needed for a refined mesh. When J gets to large it can be a good idea to instead introduce an intermediate lever between the coarse and the fine scale. The basic idea is then to solve each of the subgrid problems using the same multiscale method.

Remark 6.1 There are two constants in Theorem 6.1, $C_{inv,a}$ and C_{PF}. The first constant is the result of the bound $\sum_{i \in \mathcal{N}_0} \|\phi_i\|^2 \leq C_{inv,a}^2 \max_{\tau \in \mathcal{T}_0} h_{0,\tau}^{-2}$. We note that the left hand side is directly computable on a given coarse mesh. Furthermore, by direct computation we have, $C_{inv,a}^2 \leq 2d\gamma^2 \int_\Omega a\, dx$, where $2d$ can be replaced by $d+1$ for simplectical elements, ($2d$ is sharp for quadrilateral elements) and γ is the bound in condition of the shape regularity of the mesh.

The second constant comes from the Poincare-Freidrich type inequality $\|u_{l,J,i}\|_{L^2(\Omega)} \leq \frac{C_{PF}}{\sqrt{a_0}} \|u_{l,J,i}\|$. The constant depends on the domain Ω and can be bounded by the diameter of Ω on the convex domains considered in the numerical examples section. One could make this bound sharper by instead using an interpolation bound for $\|u_{l,J,i}\|_{L^2(\Omega)} = \|(1-\pi_0)u_{l,J,i}\|_{L^2(\Omega)}$. This would lead to an extra h_0 and also dependence of h_0/h_J. For example in two spatial dimensions we would get $\|u_{l,J,i}\|_{L^2(\Omega)} \leq C_{int} h_0 \log(h_0/h_J)^{1/2} a_0^{-1/2} \|u_{l,J,i}\|$ for some interpolation constant C_{int}, see equation (5.4) in [17].

Remark 6.2 The final estimates in Theorem 6.1 and Corollary 6.1 includes a max norm of u_J. For an appropriate triangulation (delaunay) and right hand side (bounded) it can be shown that this quantity stays bounded (independent of the mesh size) in two spatial dimensions. It can in some cases also be achieved in three spatial dimensions. However, for simplectic elements it is much more complicated than in the two dimensional case. Tetrahedrons need to have acute angles for the stiffness matrix to be an M-matrix which gives the desired max-norm bound, see e.g. [12] and references therein. The worst case scenario is that we only have the discrete Sobolev inequality $\|u_J\|_{L^\infty(\Omega)} \leq$

$C_\mathrm{d}(h_J)\|\nabla u_J\|_{L^2(\Omega)} \le C_\mathrm{d}(h_J)a_0^{-1/2}\|u_J\|$, where $C_\mathrm{d}(h_J) \sim |\log(h_J)|^{1/2}$ in two spatial dimensions and $C_\mathrm{d}(h_J) \sim h_J^{-1/2}$ in three spatial dimensions, see [3].

Remark 6.3 One can easily consider different sizes of the subdomains ω_i^k in the individual local problems. Different sizes can be handled directly since the final error bound presented in Theorem 6.1 is based on bounds for the individual basis functions $T_J\phi_i - T_J^k\phi_i$. The only difference in the final error estimate would be that the minimal k have to be used in the bound in Theorem 6.1 and Corollary 6.1.

As for the refinement levels we would have to use an interpolation estimate since $\|T_J\phi_i - T_j^k\phi_i\| \le \|T_J\phi_i - T_j\phi_i\| + \|T_j\phi_i - T_j^k\phi_i\|$ where only the second term in the right hand side can be handled by Lemma 6.1. We have excluded this generalization in order to make the paper easier to follow.

7. Numerical examples

In the numerical examples we will let $\Omega = [0, 1] \times [0, 1]$ and use continuous piecewise bilinear basis functions on a quadrilateral grid in two spatial dimensions. We let $\mathcal{T}_0$ be a coarse mesh of 30×30 rectangles. As described in the preliminaries section we let $\mathcal{T}_1, \dots, \mathcal{T}_J$ be successive uniform refinements of $\mathcal{T}_0$.

Unless otherwise stated, we let

$$(7.1) \qquad\qquad f = \chi_{\omega^\mathrm{inj}} - \chi_{\omega^\mathrm{prod}},$$

where χ_ω is the characteristic function for the set $\omega \subset \Omega$, and $\omega^\mathrm{inj} = \{(x, y) : 0 \le x, y \le 1/60\}$, $\omega^\mathrm{prod} = \{(x, y) : 1 - 1/60 \le x, y \le 1\}$. This right hand side models an injecting and a producing well in an oil reservoir. This particular example has been chosen since porous media flow is a typical application for multiscale methods. We consider five different diffusion coefficients, see Figure 2 and the definition below,

$$(7.2) \qquad \begin{cases} a_1(x, y) = 1, \\ a_2(x, y) = 1 + 0.5 \cdot sin(8x)sin(8y), \\ a_3(x, y) = 0.1 + 0.9 * \mathtt{rand}, \quad (x, y) \in \tau, \text{ for all } \tau \in \mathcal{T}_1, \\ a_4(x, y) = a_\mathrm{GSLIB}(i, j), \text{ for } \frac{i-1}{120} \le x < \frac{i}{120},\ \frac{j-1}{120} \le y < \frac{j}{120}, \\ a_5(x, y) = a_\mathrm{SPE}(i, j), \text{ for } \frac{i-1}{120} \le x < \frac{i}{120},\ \frac{j-1}{120} \le y < \frac{j}{120}, \end{cases}$$

where $\mathtt{rand}$ refers to realizations of a MATLAB function (one realization for each element in $\mathcal{T}_1$), a_GSLIB is a 120×120 matrix realization generated using GSLIB algorithms [5] and finally a_SPE is also a 120×120 matrix with entries from part of the the top layer in the tenth SPE comparative solution project (60×120 entries are used i.e. pairs of neighboring elements in the y-direction share the same value), see http://www.spe.org/web/csp/.
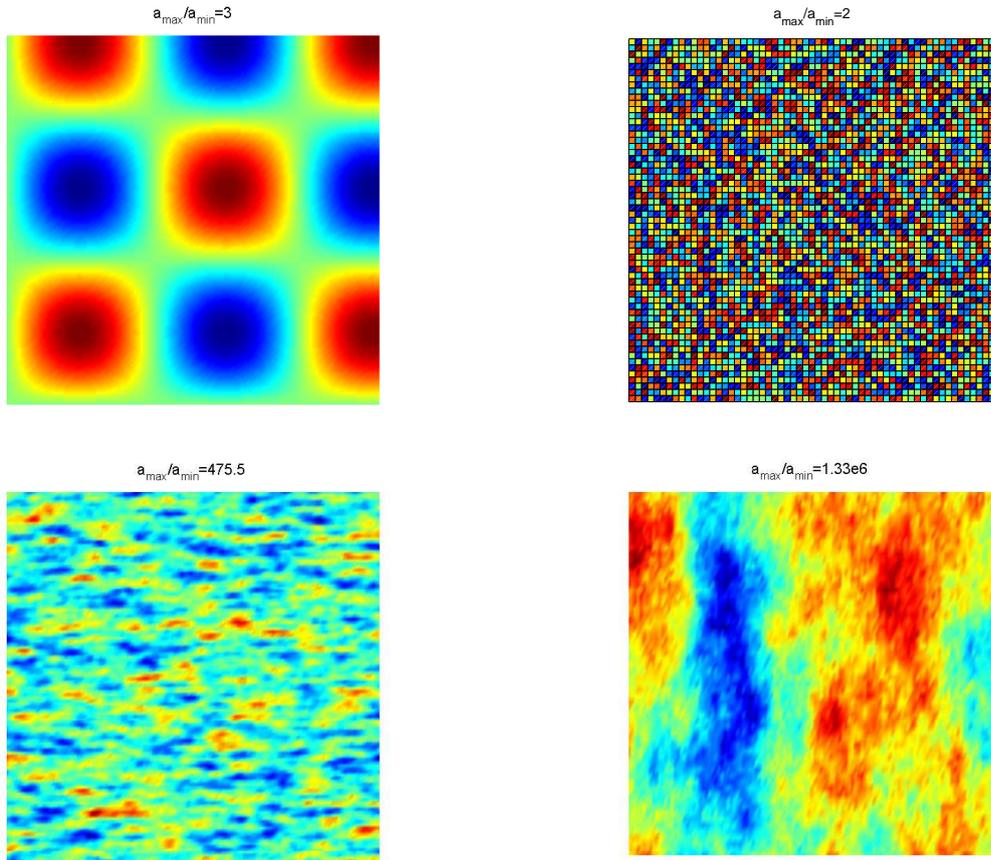
FIGURE 2. The diffusion coefficients a_i used in this paper, $i = 2, 3, 4, 5$. We skip $a_1 = 1$ and plot a_2 (left) and a_3 (right) above and the logarithm of a_4 (left) and a_5 (right) below. We also give the ratio between the maximum and the minimum value in each case.

7.1. **Convergence of the localized sub-grid solutions.** We first consider the convergence of the local subgrid problems as k (layers in the k-rings) increases. We fix $J = 3$ and k from one to five. For $a = a_5$ we plot a zoom of $T_3^{10}\phi_{435}$ in Figure 3, which is located in the center of the computational domain. Next we plot the relative error in energy norm of modified basis function number 435, i.e. $\|T_3\phi_{435} - T_3^k\phi_{435}\|/\|T_3\phi_{435}\|$ in Figure 4 (left). We also compute the corresponding errors for the same problem using the conjugated gradient method with $2k$ iterations, which is used in the proof of Lemma 6.1 and corresponds to a spread of k layers of coarse elements due to the hierarchical basis used. The result can be found in Figure 4 (right). We note that the error in the modified basis functions $(T_3\phi_i - T_3^k\phi_i)$ is not at all sensitive to the coefficient a_i while the conjugated gradient version depends directly on the condition number, and therefore on a_i. It is especially clear for a_4 and a_5. When comparing the error in the modified basis functions using $J = 2, 3$ to $J = 1$
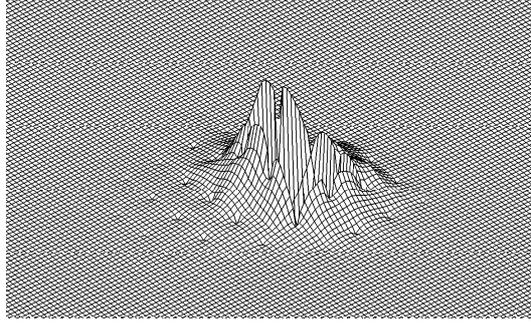
FIGURE 3. We see clearly that the function $T_3^{10}\phi_{435}$ is in the space $\mathcal{W}_J$ since it is equal to zero on the coarse nodes $\mathcal{N}_0$. We also note a very rapid decay in magnitude away form the support of the basis function ϕ_{435}.
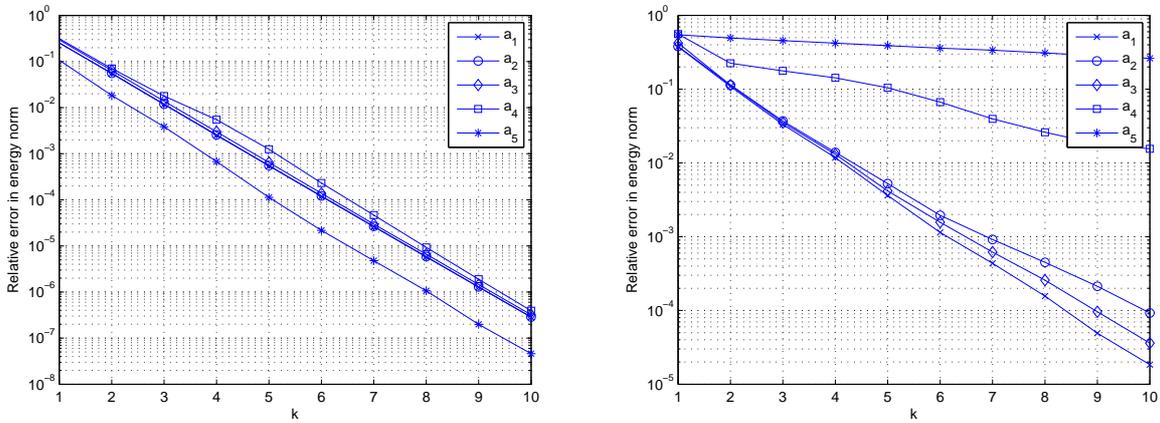

FIGURE 4. To the left we see convergence in (relative) energy norm of the error in the local basis function $\|T_3\phi_{435} - T_3^k\phi_{435}\|/\|T_3\phi_{435}\|$. To the right we instead compared the exact basis function $T_3\phi_{435}$ with $2k$ iterations of the conjugated gradient method (which gives an approximation with the same support as $T_3^k\phi_{435}$).

it seems that the convergence rate ρ depends on $(c\log(h_0/h_J) - 1)/(c\log(h_0/h_J) + 1))$ for some number $c \approx 1$, i.e. it depends on J but not on a. This positive effect can not be seen in the error analysis.

Since the bound in the estimate uses that the error in the modified basis functions can be bounded by a constant times the error in the conjugated gradient approximation it is clear that the bound will overestimate the error when the condition number is large. We also demonstrate this in Table 1, where we study the error reduction rate ρ^2 for the five different diffusion coefficients and for $J = 1, 2, 3$. Note that for $\rho^2 = 0.63$ five layers is needed to get $\rho^{2k} \leq 0.1$. The corresponding number for ten layers is 0.79. It is clear that the dependency

	a_1	a_2	a_3	a_4	a_5
$J = 1$	0.1717	0.3699	0.2430	0.7146	0.9963
$J = 2$	0.5008	0.6645	0.5814	0.9024	0.9985
$J = 3$	0.7402	0.8359	0.7887	0.9610	0.9994

TABLE 1. We compute ρ^2 for the five different diffusions and three different resolutions.
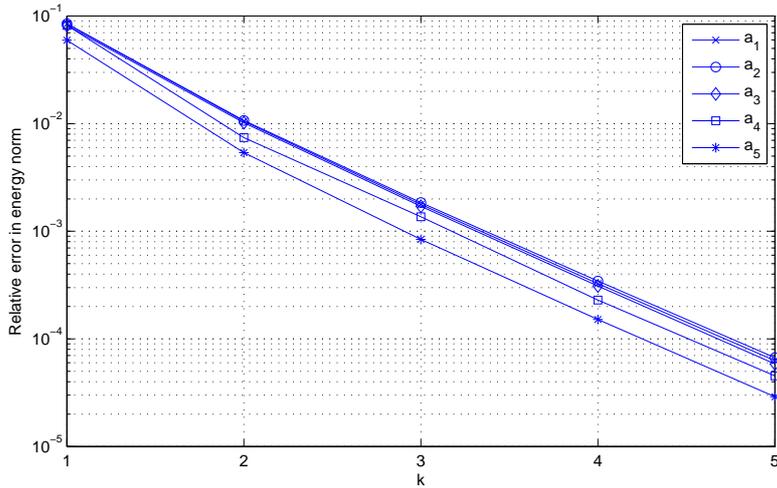

FIGURE 5. We vary k from one to five and plot the relative error in energy norm for the global solution. We have chosen $\Omega = [0, 1] \times [0, 1]$ and a 30×30 element rectangular grid as coarse mesh with a reference mesh that is uniformly refined three times, i.e. $J = 3$.

of the condition number is exaggerated in the bound (which was clear already in Figure 4) and that the analysis only gives a sensible estimate for the first three cases. On the other hand one can not expect a priori bounds to give very sharp constants. An example is the constant $C_{a,\Omega}$ in Corollary 6.1 which typically overestimates the discretization error by a large factor. In this case for a fix h_0 and J we have proven that $\|T_J \phi_i - T_J^k \phi_i\| \le C \rho^{2k}$ and the numerics confirms this, even though $0 \le \rho < 1$ gets far to close to one for cases a_4 and a_5.

7.2. **Convergence of global multiscale solution.** We are ready to study the convergence of the global multiscale solution to the reference solution in the five cases we consider. We still let $\Omega = [0, 1] \times [0, 1]$ and use a uniform rectangular grid of 30×30 elements in $\mathcal{T}_0$, and a and f given by equations (7.2) and (7.1). We fix $J = 3$ and vary the number of layers k form one to five. We plot the relative error in energy norm in Figure 5. We see that the very clear exponential decay is inherited from the modified multiscale basis functions to the global multiscale solution. Note that all errors are compared with the corresponding reference solution. So even though the error for the more complicated coefficients (a_4 and
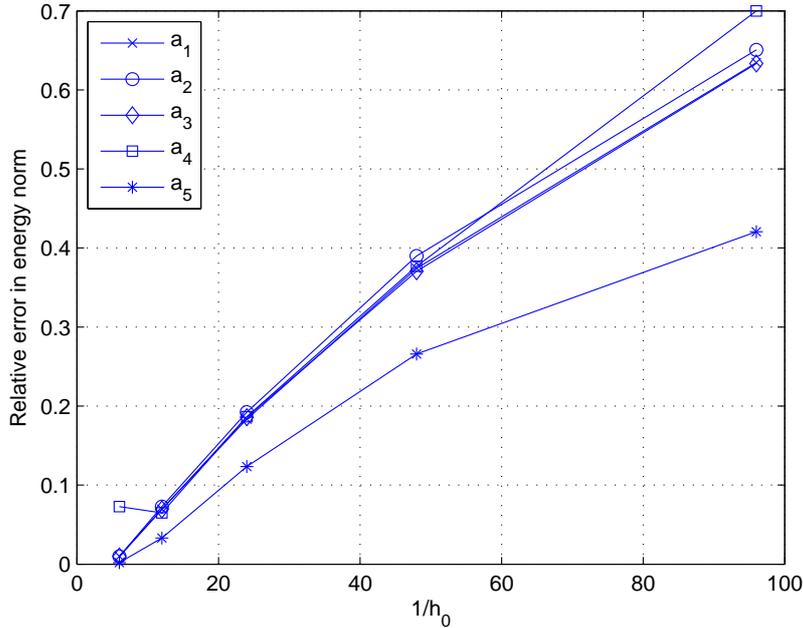
FIGURE 6. We let h_0 vary between $1/6$ and $1/96$ and let $J = 2$ and $k = 3$. We plot the relative error in energy norm.

a_5) appears to be small one must take into account that the reference solutions u_J have a larger error compared to the exact solution u for these cases. However, the method is designed to approximate the reference solution so this is the relevant error to study.

We have now studied the dependency of k for a given h_0 and J. We note that we have an explicit h_0 dependency in the bound presented in Theorem 6.1. In our final example we focus on this. We let $f = 1$ since we want to study a range of meshes and we need the right hand side to be easy to resolve. We pick $J = 2$, $k = 3$, and let $h_0 := \frac{1}{3 \cdot 2^i}$ for $i = 1, \ldots, 5$. We plot the relative error in energy norm in Figure 6. We see that the error increases when h_0 decreases and that the rate is smaller than the one presented in the bound. It is not clear if the bound is sharp in the dependency of h_0.

8. Conclusions and future work

The main result of this paper is the proof of an a priori error bound for the adaptive variational multiscale method presented in [13, 14]. The bound reveals that for a fixed coarse and reference mesh we have exponential decay of the global error in terms of the size of the local sub grid problems. This crucial result has been experimentally verified on several occasions, see e.g. [13, 14, 15, 16], but never before proven. The technique used in the proof is to bound the error in the modified basis functions by using a method where a hierarchical basis is combined with the conjugated gradient method. The bound of the error in the basis functions is then used to bound the global error. The experiments

presented in the paper is in agreement with the analysis in terms of the dependency of the crucial parameters, coarse mesh, reference mesh, and size of the local subgrid problems. The bound gives a clear over estimate of the error but this is typical for a priori error bounds of finite element methods.

However, it is also clear from the numerical experiments that the bound fails to give a good measure of the error for diffusion coefficients with high ratio $\max a / \min a$. This limitation has to do with the conjugated gradient approach used in the proof. In order to resolve this issue a more careful error analysis need to be done. On suggestion would be to use a wavelet basis instead of the hierarchical basis to reduce the error reduction rate in each iteration. We might also consider other numerical schemes instead of the conjugated gradient method to be able to prove a quicker convergence. For three dimensional problems we expect to get the same kind of convergence for fixed h_0 and J but we might get slower convergence when J increases considering equation (2.3). A natural next step would be to prove convergence for the adaptive algorithm, governing the local resolution J and subdomain size k, presented in [14].

References

[1] R. A. Adams, *Sobolev spaces,* volume 65 of Pure and Applied Mathematics, Academic Press, New York, 1975.

[2] T. Arbogast and S. L. Bryant, A two-scale numerical subgrid technique for waterflood simulations, SPE J., 446–457, Dec. 2002.

[3] J. Bramble and J. Xu, *Some estimates for a weighted L^2-projection,* Math. Comp., 56, 194, (1991), 463–476.

[4] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods,* Springer Verlag, 1994.

[5] C. V. Deutsch and A. G. Journel, *GSLIB: Geostatistical software library and user's guide, 2nd edition,* Oxford University Press, New York, 1998.

[6] J. Chu, Y. Efendiev, V. Ginting, and T. Y. Hou, *Flow based oversampling technique for multiscale finite element methods,* Advances in Water Resources, 31, (2008), 599–608.

[7] G. Golub and C. Van Loan, *Matrix computations,* 3rd edition, John Hopkins University Press, Baltimore, MD, 1996.

[8] 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, (1997), 169–189.

[9] T. Y. Hou and X.-H. Wu, and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients,* Math. Comp., 68, 227, (1999), 913–943.

[10] T. J. R. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods,* Comput. Methods Appl. Mech. Engrg., 127, (1995), 387–401.

[11] T. 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 (1998) 3–24.

[12] J. Karátson, S. Korotov, and M. Krížek, *On discrete maximum principles for nonlinear elliptic problems,* Math. Comput. Simulation 76 (2007), no. 1-3, 99-108.

[13] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori error estimation: Duality techniques for elliptic problems,* Lecture Notes in Computational Science and Engineering, B. Engquist, P. Lötstedt, O. Runborg (eds.) Vol. 44, 181–193, 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 (2007) 2313-2324.

[15] M. G. Larson and A. Målqvist, *Adaptive variational multiscale method of convection-diffusion problems,* Comm. Num. Methods Engrg., 25, (2009), 65–79.

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

[17] M. Marion and J. Xu, *Error estimates on a new nonlinear Galerkin method based on two grid finite elements,* SIAM J. Numer. Anal., 32, (1995), 1170–1184.

[18] P. Morin, R. H. Nochetto, and K. G. Siebert, *Local problems on stars: A posteriori error estimators, convergence, and performance,* Math. Comp., 72, (2003), 1067–1097.

[19] J. Nolen, G. Papanicolaou, and O. Pironneau, *A framework for adaptive multiscale methods for elliptic problems,* Multiscale Model. Simul., 7, (2008), no. 1, 171–196.

[20] J. M. Nordbotten, *Adaptive variational multiscale methods for multi-phase flow in porous media,* Multiscale Model. Simul., 7, (2008), 1455–1473.

[21] H. Schroll and A. Tveito, *Local existence and stability for a hyperbolic-elliptic system modeling two-phase reservoir flow,* Electron. J. Differential Equations 2000, No. 04, 28 pp.

[22] J. Xu, *Iterative methods by space decomposition and subspace correction,* SIAM Review, 34, (1992), 581–613.

DEPARTMENT OF INFORMATION TECHNOLOGY, UPPSALA UNIVERSITY, UPPSALA, SWEDEN
E-mail address: axel.malqvist@it.uu.se
URL: http://user.it.uu.se/~axelm