Multiscale Model. Simul.
Vol. 9, No. 3, pp. 1064–1086

MULTISCALE METHODS FOR ELLIPTIC PROBLEMS*

AXEL MÅLQVIST†

Abstract. In this paper we derive a framework for multiscale approximation of elliptic problems on standard and mixed form. The method presented is based on a splitting into coarse and fine scales together with a systematic technique for approximation of the fine scale part, based on the solution of decoupled localized subgrid problems. The fine scale approximation is then used to modify the coarse scale equations. A key feature of the method is that symmetry in the bilinear form is preserved in the discrete system. Other key features are a posteriori error bounds and adaptive algorithms based on these bounds. The adaptive algorithms are used for automatic tuning of the method parameters. In the last part of the paper we present numerical examples where we apply the framework to a problem in oil reservoir simulation.

Key words. multiscale methods, error analysis, adaptivity, mixed methods, finite element method

AMS subject classifications. 65N30, 65N15, 65N50

DOI. 10.1137/090775592

1. Introduction.

There are numerous multiscale problems in engineering applications. They appear in all branches of the engineering sciences, for instance, composite materials, flow in porous media, and fluid mechanics. A common feature for all these applications is that they are very computationally challenging and often impossible to solve within an acceptable tolerance using standard one mesh methods. We thus need to develop new methods that are based on a combination of global and local computations: so called multiscale methods. Furthermore, to guarantee accuracy we need to derive error bounds and based on these bounds, adaptive algorithms for automatic tuning of the method parameters.

Previous work. Various multiscale methods have been developed during the last two decades. In [9], Hou and Wu present the multiscale finite element method. The idea is based on homogenization theory. Decoupled localized fine scale problems are solved on each coarse element, using a finer local mesh, in order to modify the coarse scale basis functions. The coarse scale equations are then solved using 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 patches, called oversampling, has been introduced; see, e.g., [7]. The variational multiscale method (VMS) is an alternative approach that serves as a general framework for solving multiscale problems; 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].

The adaptive variational multiscale method (AVMS) was introduced by Larson and Målqvist in [12]. The development was inspired by the solution of local problems on stars

*Received by the editors November 1, 2009; accepted for publication (in revised form) April 19, 2011; published electronically July 28, 2011.

http://www.siam.org/journals/mms/9-3/77559.html

†Department of Information Technology, Uppsala University, SE-751 05, Uppsala, Sweden. (axel.malqvist@it.uu.se). The author's work is supported by The Göran Gustafsson Foundation.

used in [16] 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 patches that may be larger than the single mesh stars used in [16]. The paper [12] presents a posteriori error estimates where the error is bounded in terms of fine mesh size, coarse mesh size, and size of the local subdomains. The method has been further developed in several directions: to convection diffusion problems [14] with error estimates for linear functionals, to parabolic problems [19], and to a mixed formulation of the Poisson equation [15]. In [15], symmetry of the bilinear form was not preserved, but a symmetric version was briefly discussed in Remark 2.1. Another recent work is [18], where a very similar method is presented.

New contributions. In this paper we present an abstract framework for constructing multiscale methods for elliptic partial differential equations. We consider both symmetric and nonsymmetric problems and derive general methods that preserve symmetry if the original problem is symmetric. We derive a posteriori error estimates in global norms for two specific problems, a convection-diffusion-reaction equation and the Poisson equation on mixed form. The error estimate for the mixed problem is an extension of the result presented in Corollary 3.1 in [15] to a symmetric formulation. The estimate for the convection-diffusion-reaction problem is new. Based on the error estimates, we formulate adaptive algorithms that automatically tune the method parameters according to computable error indicators. The adaptive nature of the method is a crucial feature since multiscale problems are very computationally challenging. We apply the adaptive method to an oil reservoir problem and obtain very promising numerical results.

Outline. In section 2 we present an abstract framework for constructing multiscale methods. In section 3 we consider a convection-diffusion-reaction equation and derive a posteriori error estimates. In section 4 we consider the Poisson equation on mixed form and present error estimates. In section 5 we present an adaptive strategy, and in section 6 we give numerical examples. In section 7 we present conclusions, and, finally, in section 8 we present proofs of the theoretical results.

2. Framework for constructing multiscale methods. We consider a second order elliptic partial differential equation with multiscale coefficients, on weak form. We present a technique for constructing efficient methods for computing the solution using a coarse scale, where the coefficients are not resolved but a system of equations can be handled on a single mesh, and a fine scale where the coefficients are resolved, but where the linear systems are huge and need to be localized and handled in a parallel fashion. We start by presenting some notations.

2.1. Function spaces. Let $\omega \subset \mathbf{R}^d$, $d = 2, 3$ be a domain, with Lipschitz boundary $\partial\omega$. The outward normal of the boundary is denoted n. We let $W^{k,p}(\omega)$ be the Sobolev space of degree p and smoothness k with corresponding norm $\|\cdot\|_{W^{k,p}(\omega)}$. We adopt the simplified notations $H^k(\omega) = W^{k,2}(\omega)$ and $L^p(\omega) = W^{0,p}(\omega)$. We will also use the spaces $H(\mathrm{div};\Omega) = \{v \in L^2(\omega) : \nabla \cdot v \in L^2(\omega)\}$ and $H_0^1(\omega) = \{v \in H^1(\omega) : v|_{\partial\omega} = 0\}$. The restriction to the boundary $v_{|\partial\omega}$ is interpreted in the sense of traces. The dual of $H_0^1(\omega)$ will be denoted $H^{-1}(\omega)$. Further, we let $(v, w)_\omega = \int_\omega vw\mathrm{d}x$ with short notation $(v, w) = (v, w)_\Omega$, and $\langle v, w\rangle_{\partial\omega} = \int_{\partial\omega} vw\mathrm{d}x$. More information on these function spaces can be found in [1]. We will not track constants in this paper; i.e., a constant C can mean different things in different equations. In particular we treat $\|a \cdot\|_{L^2(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$ as equivalent norms when a is bounded from above and below by positive numbers. The effect of a will be hidden in the constants C.

2.2. Abstract model problem on weak form. We study an elliptic equation with variable coefficient and homogeneous boundary conditions. Let $\mathcal{V}$ be a Sobolev space defined on a Lipschitz domain $\Omega \subset \mathbf{R}^d$ with boundary $\partial\Omega$. We let $\|\cdot\|_{\mathcal{V}}$ be a norm on $\mathcal{V}$. Further, let $a\colon \mathcal{V} \times \mathcal{V} \to \mathbf{R}$ be a bilinear form and $l\colon \mathcal{V} \to \mathbf{R}$ be a linear functional. We seek a solution to the following variational problem: find $u \in \mathcal{V}$ such that

$$a(u, v) = l(v) \quad \text{for all } v \in \mathcal{V}. \tag{2.1}$$

We assume that there exists a unique solution to (2.1). Before we proceed with the variational multiscale framework we present two problems that fit this framework.

Convection-diffusion-reaction equation. We let $\mathcal{V} = H_0^1(\Omega)$ and define $\|\cdot\|_{\mathcal{V}} = \|\cdot\|_{H^1(\Omega)}$. The linear functional l and the bilinear form a are defined by

$$l(v) = (f, v) \qquad \text{for all } v \in \mathcal{V}, \tag{2.2}$$

$$a(v, w) = (a\nabla v, \nabla w) + (b \cdot \nabla v, w) + (cv, w) \qquad \text{for all } v, w \in \mathcal{V}. \tag{2.3}$$

Poisson equation on mixed form. We let $\mathcal{V} = L^2(\Omega)/\mathbf{R} \times \{v \in H(\mathrm{div}; \Omega)\colon n \cdot v = 0\}$ and define $\|\cdot\|_{\mathcal{V}} = \|\frac{1}{\sqrt{a}} \cdot \|_{L^2(\Omega)}$. The linear functional l and the bilinear form a are defined by

$$l(w) = (f, w_1) \qquad \text{for all } w \in \mathcal{V}, \tag{2.4}$$

$$a(v, w) = \left(\frac{1}{a} v_2, w_2\right) + (v_1, \nabla \cdot w_2) - (\nabla \cdot v_2, w_1) \qquad \text{for all } v, w \in \mathcal{V}, \tag{2.5}$$

where subscript 1 refers to the component in $L^2(\Omega)/\mathbf{R}$ and 2 refers to the component in $\{v \in H(\mathrm{div}; \Omega)\colon n \cdot v = 0\}$.

We will return to these two examples later in the paper, but now we continue with the general framework.

2.3. Split between coarse and fine scale. Let $\mathcal{V}_c \subset \mathcal{V}$ be a finite element space defined on a fix given coarse mesh $\mathcal{K}_H = \{K\}$ of shape regular elements such that $\cup K = \Omega$ and K are disjoint. Let $H_K = \mathrm{diam}(K)$ be the mesh parameter and let $H = \max_{K \in \mathcal{K}_H} H_K$. Further assume that there exists a positive number α such that $\alpha^{-1}\|v\|_{\mathcal{V}}^2 \leq a(v, v)$ for all $v \in \mathcal{V}$. In standard finite element methods for solving elliptic partial differential equations, the Ritz projection $\mathcal{R}_c\colon \mathcal{V} \to \mathcal{V}_c$, defined by $a(\mathcal{R}_c u, v_c) = a(u, v_c)$ for all $v_c \in \mathcal{V}_c$, of the exact solution onto $\mathcal{V}_c$ is computed. In a multiscale setting it is known that $\|(I - \mathcal{R}_c)u\|_{\mathcal{V}}$ is not necessarily small if the fine scales are not resolved. If we consider the Poisson equations with periodic diffusion coefficient (period ϵ) solved using linear finite elements on a mesh with mesh size H, we have

$$\|(I - \mathcal{R}_c)u\|_{H^1(\Omega)} \leq C\frac{H}{\epsilon} \tag{2.6}$$

for some constant $C > 0$; see, e.g., [9]. By using the Ritz projection the fine scale features of the solution are neglected. In VMS a split between coarse and fine scales is introduced using an inclusion operator $\mathcal{I}_c\colon \mathcal{V} \to \mathcal{V}_c$. For example, $\mathcal{I}_c$ can be the nodal interpolant or the L^2 projection onto $\mathcal{V}_c$. An approximation $u_c + u_f$ is sought in $\mathcal{V}_c \bigoplus \mathcal{V}_f$, where $\mathcal{V}_c = \mathcal{I}_c(\mathcal{V})$ and $\mathcal{V}_f = (I - \mathcal{I}_c)\mathcal{V} = \{v \in \mathcal{V}\colon \mathcal{I}_c v = 0\}$. This results in a set of coarse scale equations and a set of fine scale equations: find $u_c \in \mathcal{V}_c$ and $u_f \in \mathcal{V}_f$ such that

$$a(u_c, v_c) + a(u_f, v_c) = l(v_c) \qquad \text{for all } v_c \in \mathcal{V}_c, \tag{2.7}$$

$$a(u_f, v_f) = l(v_f) - a(u_c, v_f) \qquad \text{for all } v_f \in \mathcal{V}_f. \tag{2.8}$$

We note that if the fine scales are neglected ($u_f = 0$), we get back the Ritz projection on the coarse mesh.

In this paper we will consider a different split. Our goal is to isolate the fine scales and thus get coarse and fine scale equations that can be solved independently. In particular if a good approximation of $\mathcal{I}_c u$ is sought, only the coarse scale equation should have to be solved. We introduce the following multiscale projection operator $\mathcal{T}: \mathcal{V}_c \to \mathcal{V}_f$:

$$a(\mathcal{T} v_c, v_f) = -a(v_c, v_f) \qquad \text{for all } v_c \in \mathcal{V}_c \quad \text{and} \quad v_f \in \mathcal{V}_f. \tag{2.9}$$

Since we consider nonsymmetric bilinear forms, we also introduce $\mathcal{T}^*: \mathcal{V}_c \to \mathcal{V}_f$ defined by

$$a(v_f, \mathcal{T}^* v_c,) = -a(v_f, v_c) \qquad \text{for all } v_c \in \mathcal{V}_c \quad \text{and} \quad v_f \in \mathcal{V}_f. \tag{2.10}$$

The existence of solutions $\mathcal{T} v_c$ and $\mathcal{T}^* v_c$ for any given $v_c \in \mathcal{V}_c$ follows from the assumptions on a guaranteeing existence and uniqueness for the original problem (2.1). By expanding u and v in (2.1) using $\mathcal{T}$ and $\mathcal{T}^*$—i.e., $u = u_c + \mathcal{T} u_c + u_f$ and $v = v_c + \mathcal{T}^* v_c + v_f$—we get the following linear system of equations: find $u_c \in \mathcal{V}_c$ and $u_f \in \mathcal{V}_f$ such that

$$a(u_c + \mathcal{T} u_c, v_c + \mathcal{T}^* v_c) = l(v_c + \mathcal{T}^* v_c) \qquad \text{for all } v_c \in \mathcal{V}_c, \tag{2.11}$$

$$a(u_f, v_f) = l(v_f) \qquad \text{for all } v_f \in \mathcal{V}_f. \tag{2.12}$$

Note that u_f plays a different role here than in (2.7) and (2.8), since also $\mathcal{T} u_c \in \mathcal{V}_f$. We need to approximate $\mathcal{T}$ and $\mathcal{T}^*$ in order to compute the coarse scale solution. If we want the fine scale part $(1 - \mathcal{I}_c)u$, we also need to approximate u_f. This means inverting linear systems on the fine scale, which appears to be just as expensive as solving the original problem. However, there is a difference. Test and trail space only include the fine scale features of the problem; i.e., the space in which we seek a solution is $\mathcal{V}_f$ instead of $\mathcal{V}$. It turns out that the behavior of the bilinear form is much more local in $\mathcal{V}_f$ than in $\mathcal{V}$. This means that $\mathcal{T} v_c$, $\mathcal{T}^* v_c$, and u_f can be computed efficiently by localizing the right-hand sides in space and solving the remaining problems independently, in parallel, on small overlapping patches with homogeneous boundary conditions. We will illustrate the quick decay in $\mathcal{V}_f$ by solving a variational problem with bilinear form a and a localized right-hand side.

2.4. Example of decay in $\mathcal{V}_f$ compared to $\mathcal{V}$. Let $\mathcal{K}_H$ be a coarse mesh of 8×8 square elements defined on a unit square Ω and let $\mathcal{V}_c \subset H_0^1(\Omega)$ be the corresponding finite element space of continuous piecewise bilinear functions. In particular let $\varphi_i \in \mathcal{V}_c$ be a basis function equal to one in $(3/8, 3/8)$ and zero in all other nodes. We consider the problem, find u such that

$$-\Delta u = \varphi_i \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega. \tag{2.13}$$

We note that φ_i has local support in Ω. The weak form reads as follows: find $u \in \mathcal{W}$ such that

$$(\nabla u, \nabla v) = (\varphi_i, v) \quad \text{for all } v \in \mathcal{W}, \tag{2.14}$$

where $\mathcal{W} = H_0^1(\Omega)$. We now want to show that if we solve (2.14) in a slice space $\mathcal{W} = \mathcal{V}_f$, where coarse functions are left out, the decay of the solution u will be very quick. This makes it possible to get a good approximation by solving the problem on a subdomain $\omega \subset \Omega$ such that $\text{supp}(\varphi_i) \subset \omega$ using homogeneous Dirichlet boundary conditions on $\partial\omega$. The shape of these subdomains is arbitrary, but in practice we will focus on a simple constructing that takes advantage of the data structures available for the coarse mesh.

DEFINITION 2.1. *We say that ω_i^1 is a 1-layer patch if $\omega_i^1 = \text{supp}(\phi_i)$, where ϕ_i is a coarse basis function. Further we say that ω_i^n is an n-layer patch if*

$$\omega_i^n = \cup_{\{i:\text{supp}(\phi_i)\cap\omega_i^{n-1}\neq\emptyset\}} \text{supp}(\phi_i), \qquad n = 2, 3, \ldots \tag{2.15}$$

We sometimes omit the superscript n if it is arbitrary or given from the context.

This means that the one layer patch is the support of the basis function, and for each layer a ring of coarse elements is added to the patch; see Figure 2.1 for an illustration of a one-layer and a two-layer patch using a Lagrangian coarse basis function ϕ_i.

We consider the space $\mathcal{V} = H_0^1(\omega_i^3)$, where ω_i^3 is a three layer patch of coarse elements surrounding $(3/8, 3/8)$; i.e., $\omega_i^3 = [0, 3/4] \times [0, 3/4]$. We further let π_c be the nodal interpolant and P_c be the $L^2(\omega_i^3)$ projection onto $\{v \in \mathcal{V}_c : \text{supp}(v) \subset \omega_i^3\}$. We use truncated Fourier series expansion to compute solutions u. First we seek a solution in the entire space $\mathcal{W} = \mathcal{V}$, then we let $\mathcal{W} = (I - \pi_c)\mathcal{V}$, and finally we let $\mathcal{W} = (I - P_c)\mathcal{V}$. The constraints are realized using Lagrange multipliers. We present the result in Figure 2.2. We note that the problems solved in the fine scale slice spaces show very rapid decay compared to the problem solved in the entire space $\mathcal{V}$. In Figure 2.3 we vary the size of the patch $\omega_i^L \subset \Omega$ from $L = 1$ to $L = 4$ (see Definition 2.1) and plot the size of the normal derivative integrated over the boundary $|\int_{\partial\omega_i^L} \partial_n u \, ds|$. This quantity indicates how big the effect of increasing the patch size is. We see exponential decay with respect to distance measured in the number of layers of the patch for the problems solved in the slice spaces $\mathcal{V}_f$. For the solution in the entire space $H_0^1(\omega_i^L)$, we

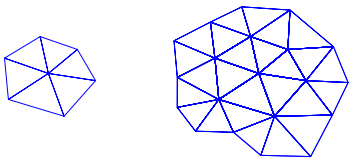

FIG. 2.1. *Examples of one (left) and two (right) layer patches on a triangular mesh.*

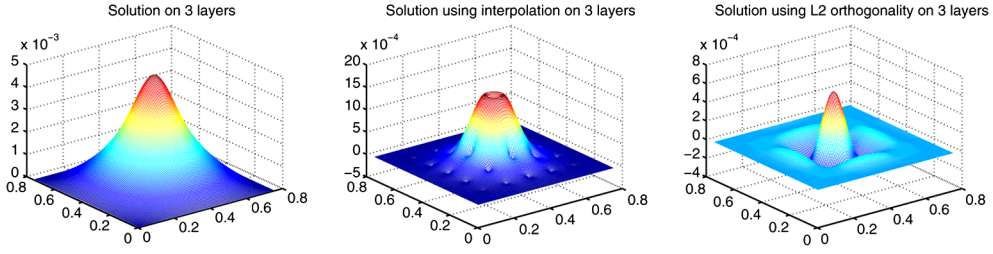

FIG. 2.2. *Equation (2.14) solved in different spaces $\mathcal{W}$. To the left $\mathcal{W} = \mathcal{V} = H_0^1(\omega_i^3)$, middle $\mathcal{W} = (I - \pi_c)\mathcal{V}$ using hierarchical split, and right $\mathcal{W} = (I - P_c)\mathcal{V}$ using L^2-orthogonal split.*

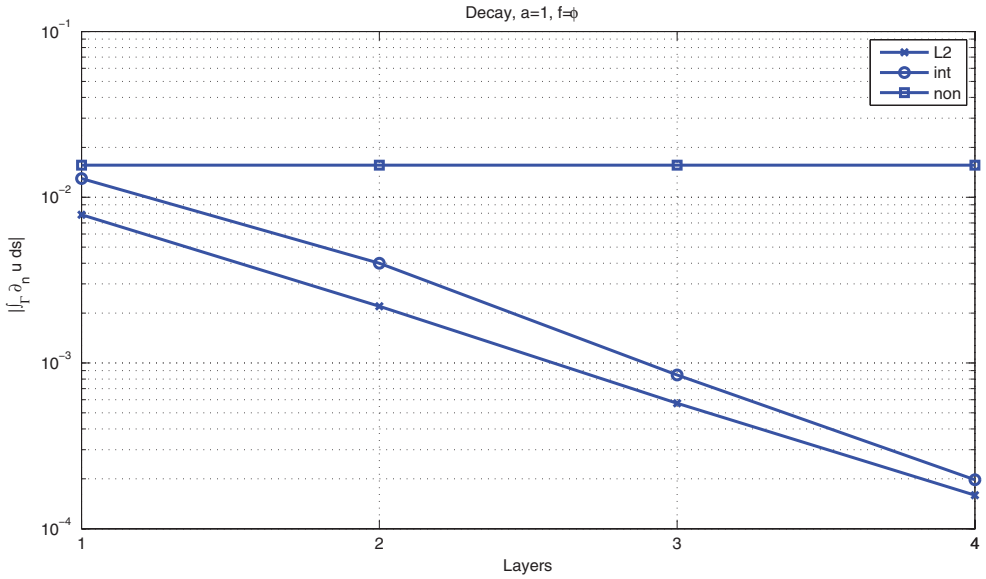

FIG. 2.3. *Decay of* $|\int_{\partial\omega_i^t} \partial_n u \mathrm{d}s|$ *as a function of layers in the patch in for whole space (non) fine part using the nodal interpolant as splitting operator (int) and fine part using* L^2 *projection as splitting operator (L2).*

get a value independent of patch size, as expected since all production is within the one layer patch.

This observation leads us to the idea of localizing the calculations of $\mathcal{T} v_c$, $\mathcal{T}^* v_c$, and u_f in space and then solving them independently, in parallel, using homogeneous boundary conditions on patches. If the decay is quick, the number of layers needed to get a good approximation will be small and therefore also will overlap. Later in the paper we will show how a posteriori error estimates capture the sources of error and can be used to automatically tune mesh size and the size of the patches. We will also show more examples of the decay rate for more interesting coefficients a. A priori bounds that capture the decay theoretically have not yet been proved. This is work in progress.

3. Application to convection-diffusion-reaction equations using Lagrangian elements. We will now show how the abstract framework can be used as a starting point for constructing a multiscale method for convection-diffusion-reaction equations. We use (2.2) and (2.3) as model problems and let $\|\cdot\|_{\mathcal{V}} = \|\cdot\|_{H^1(\Omega)}$. On strong form we introduce the differential operator $\mathcal{L}$ in order to simplify notations. We let

$$\mathcal{L}u := -\nabla \cdot a\nabla u + b \cdot \nabla u + cu = f \qquad \text{in } \Omega, \tag{3.1}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{3.2}$$

where $f \in L^2(\Omega)$, and a, b, $c \in L^\infty(\Omega)$ such that

$$\|v\|_{\mathcal{V}} \leq \alpha\, a(v, v) \qquad \text{for all } v \in \mathcal{V}, \tag{3.3}$$

$$a(v, w) \leq \beta \|v\|_{\mathcal{V}} \|w\|_{\mathcal{V}} \qquad \text{for all } v, w \in \mathcal{V}, \tag{3.4}$$

$$l(v) \leq \gamma \|v\|_{\mathcal{V}} \qquad \text{for all } v \in \mathcal{V} \tag{3.5}$$

for positive constants α, β, $\gamma \in \mathbf{R}$. This implicitly puts restrictions on the coefficients. The diffusion coefficient a needs to be positive bounded away from zero, c should typically be nonnegative, and b is small enough so that $\alpha\|v\|^2_{H^1(\Omega)} \leq a(v,v)$ holds. The weak form using the operator $\mathcal{L}$ reads as follows: find $u \in \mathcal{V}$ such that

$$(\mathcal{L}u, v) = a(u,v) = l(v) = (f,v) \qquad \text{for all } v \in H^1_0(\Omega). \tag{3.6}$$

Lax–Milgram's lemma guarantees the existence of a unique solution to the weak form.

3.1. Notations. We introduce a finite element space $\mathcal{V}_c$ of continuous piecewise linear functions on the mesh $\mathcal{K}_H$ of shape regular elements. We let $\{\boldsymbol{\phi}_i\}_{i\in\mathcal{N}}$, for an index set over nodes $\mathcal{N}$, be the Lagrangian basis functions spanning the space $\mathcal{V}_c = \text{span}(\{\boldsymbol{\phi}_i\}_{i\in\mathcal{N}})$. As inclusion operator $\mathcal{I}_c : \mathcal{V} \to \mathcal{V}_c$, we choose the nodal interpolant, which means that we use an hierarchical split between the coarse and fine scales. Since subgrids will be used frequently we also introduce the notation $\mathcal{K}_H(\omega) = \{K \in \mathcal{K}_H : K \in \omega\}$, where $\omega \subset \Omega$ such that ω is a union of coarse elements K; i.e., $\cup_{K\in\mathcal{K}_H(\omega)} K = \omega$.

The fine mesh will only be defined on local subregions $\omega \subset \Omega$, made up of coarse elements. We let $\mathcal{K}_h(\omega) = \{K\}$ such that $\cup_{K\in\mathcal{K}_h(\omega)} K = \omega$ and K are disjoint and shape regular. Further we assume that the meshes are nested so that any $K \in \mathcal{K}_H(\omega)$ can be written as a union of elements in $\mathcal{K}_h(\omega)$. We let $h^\omega_K = \text{diam}(K)$ for each $K \in \mathcal{K}_h(\omega)$. We omit the superscript when it is obvious which domain we are referring to. We let $\mathcal{V}^h(\omega)$ be the space of continuous piecewise linear functions on the mesh $\mathcal{K}_h(\omega)$, with vanishing trace, and proceed to define a finite element approximation of $\mathcal{V}_f$ restricted to ω. Let

$$\mathcal{V}^h_f(\omega) = \{v \in \mathcal{V}^h(\omega) : \mathcal{I}_c v = 0\}. \tag{3.7}$$

We will denote the Scott–Zhang interpolant onto the coarse space $\pi_c : \mathcal{V} \to \mathcal{V}_c$; see [5].

3.2. The multiscale method. We need to solve three different kinds of equations on the fine scale to get hold of $\mathcal{T}v_c$, $\mathcal{T}^*v_c$, and u_f. As seen in the previous section, we want to localize the computation on the fine scale. We are going to use the coarse scale basis functions $\{\boldsymbol{\phi}_i\}_{i\in\mathcal{N}}$ to do this. We define localized approximations of the three problems. For each member of the partition of unity $\{\boldsymbol{\phi}_i\}_{i\in\mathcal{N}}$, we introduce a patch ω_i according to Definition 2.1. We get the following: find $\tilde{\mathcal{T}}_i v_c \in \mathcal{V}^h_f(\omega_i)$, $\tilde{\mathcal{T}}^*_i v_c \in \mathcal{V}^h_f(\omega_i)$, and $U_{f,i} \in \mathcal{V}^h_f(\omega_i)$ such that

$$a(\tilde{\mathcal{T}}_i v_c, v_f) = -a(v_c, \boldsymbol{\phi}_i v_f) \quad \text{for all } v_c \in \mathcal{V}_c, \quad v_f \in \mathcal{V}^h_f(\omega_i), \tag{3.8}$$

$$a(v_f, \tilde{\mathcal{T}}^*_i v_c) = -a(\boldsymbol{\phi}_i v_f, v_c) \quad \text{for all } v_c \in \mathcal{V}_c, \quad v_f \in \mathcal{V}^h_f(\omega_i), \tag{3.9}$$

$$a(U_{f,i}, v_f) = l(\boldsymbol{\phi}_i v_f) \qquad \text{for all } v_f \in \mathcal{V}^h_f(\omega_i) \tag{3.10}$$

for all $i \in \mathcal{N}$. We note that we only need to test with basis functions $\boldsymbol{\phi}_j \in \mathcal{V}_c$ that overlap $\text{supp}(\boldsymbol{\phi}_i)$, since the right-hand side otherwise is zero. We define $\tilde{\mathcal{T}} = \sum_{i\in\mathcal{N}} \tilde{\mathcal{T}}_i$, $\tilde{\mathcal{T}}^* = \sum_{i\in\mathcal{N}} \tilde{\mathcal{T}}^*_i$, and $U_f = \sum_{i\in\mathcal{N}} U_{f,i}$. The well posedness of these equations follows from the assumptions on a and l, using Lax–Milgram's lemma, if we can prove that $\|\boldsymbol{\phi}_i v_f\|_{\mathcal{V}} \leq C\|v_f\|_{\mathcal{V}}$, since this is needed to prove the boundedness of the right-hand sides. We prove this in the following lemma.

LEMMA 3.1. *For any function* $v_f \in \mathcal{V}_f$ *and all basis functions* $\{\boldsymbol{\phi}_i\}_{i\in\mathcal{N}} \in \mathcal{V}_c$, *it holds that*

$$\|\boldsymbol{\phi}_i v_f\|_{H^1(\Omega)} \leq C\|v_f\|_{H^1(\Omega)}, \tag{3.11}$$

where C *is independent of the coarse scale mesh size in three spatial dimensions and is bounded by a constant times* $(\min_{K\in\mathcal{K}_H} H_K)^{-\epsilon}$ *for any* $\epsilon > 0$ *in two spatial dimensions.*

Proof. See section 8.1. ∎

The solutions to (3.8)–(3.10) are all computable. We can immediately formulate the method on the coarse scale, using (2.11) as a starting point. The method reads as follows: find $U_c \in \mathcal{V}_c$ such that

$$a(U_c + \tilde{\mathcal{T}} U_c, v_c + \tilde{\mathcal{T}}^* v_c) = l(v_c + \tilde{\mathcal{T}}^* v_c) - a(U_f, v_c + \tilde{\mathcal{T}}^* v_c) \qquad \text{for all } v_c \in \mathcal{V}_c. \tag{3.12}$$

There is one important difference between (2.11) and (3.12). The term $a(U_f, v_c + \tilde{\mathcal{T}}^* v_c)$ is added. This term vanishes in the continuous equation since $a(u_f, v_c + \mathcal{T}^* v_c) = 0$. Without this term we would not have Galerkin orthogonality between the error $u - (U_c + \tilde{\mathcal{T}} U_c + U_f)$ and the space $\{v_c + \tilde{\mathcal{T}}^* v_c, v_c \in \mathcal{V}_c\}$, which seems unnatural for a finite element method and would make the error analysis more complicated. In the multiscale finite element method the equivalent term is left out while in the VMS it is present. However, in the VMS, $\tilde{\mathcal{T}}^* v_c$ is instead replaced by zero.

3.3. Error estimation. The next step is to bound the error between the solution to (3.1) and (3.2) u and the solution to (3.12) $U_c + \tilde{\mathcal{T}} U_c + U_f$. As a crucial step of the proof we need to make sure that $\tilde{\mathcal{T}}^*$ is stable in $H^1(\boldsymbol{\Omega})$.

LEMMA 3.2. *It holds that*

$$\|\tilde{\mathcal{T}}^* v_c\|_{H^1(\Omega)} \leq C\|v_c\|_{H^1(\Omega)} \qquad for\ all\ v_c \in \mathcal{V}_c. \tag{3.13}$$

Proof. See section 8.2. ∎

We are now ready to present an a posteriori error bound for the proposed method. We employ the notation $[\cdot]$ for the jumps of discontinuous functions over element faces.

THEOREM 3.1. *The difference between* u, *solving* (3.1) *and* (3.2), *and* $U = U_c + \sum_{i\in\mathcal{N}}(U_{f,i} + \tilde{\mathcal{T}}_i U_c)$, *where* $U_c \in \mathcal{V}_c$ *and* $U_{f,i}, \tilde{\mathcal{T}}_i U_c \in \mathcal{V}_f^h(\omega_i)$ *are defined by* (3.8)–(3.10) *and* (3.12), *can be bounded as follows:*

$$\|u - U\|_{H^1(\Omega)} \leq C\left(\sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} \rho_K^i(U)\right)^{1/2} + C\sum_{i\in\mathcal{N}} \max_{K\in\mathcal{K}_H(\omega_i)} H_K^{1/2} \|\Sigma_f^i(U)\|_{L^2(\partial\omega_i)}, \tag{3.14}$$

where

$$\begin{aligned} \rho_K^i(U) = {} & h_K^2 \|\boldsymbol{\phi}_i(f - \mathcal{L}U_c) - \mathcal{L}(U_{f,i} + \tilde{\mathcal{T}}_i U_c)\|_{L^2(K)}^2 \\ & + h_K \|[a\boldsymbol{\phi}_i \partial_n U_c + a\partial_n(U_{f,i} + \tilde{\mathcal{T}}_i U_c)]\|_{L^2(\partial K)}^2 \end{aligned} \tag{3.15}$$

for all $i \in \mathcal{N}$ *and* $\Sigma_f^i(U) \in \mathcal{V}_f^h(\partial\omega_i) = \{v \in L^2(\partial\omega_i) : v = w|_{\partial\omega_i}\ for\ some\ w \in \mathcal{V}_f^h(\omega_i)\}$ *are defined individually on each patch in the following way:*

$$\langle \Sigma_f^i(U), v_f \rangle_{\partial\omega_i} = l(\boldsymbol{\phi}_i v_f) - a(U_c, \boldsymbol{\phi}_i v_f) - a(U_{f,i} + \tilde{\mathcal{T}}_i U_c, v_f) \qquad for\ all\ v_f \in \mathcal{V}_f^h(\partial\omega_i). \tag{3.16}$$

Proof. See section 8.3 ∎

Remark 3.1. We see clearly in (3.14) that the error is bounded by two terms, where the first term depends on the fine scale mesh size and the second term depends on the decay of the function $\Sigma_f^i(U)$, which depends directly on the decay rate of the local solution $\tilde{\mathcal{T}}_i U_c + U_{f,i}$. The first term of the error bound can therefore be reduced by decreasing the fine scale mesh size (locally on each patch), while the second term is reduced by increasing the size of the patches. We view the coarse mesh with mesh size H as fixed and given in this paper.

4. Application to a mixed problem using Raviart–Thomas elements. We turn to our second model problem. We consider (2.4) and (2.5) and use the energy norm $\|v\|_{\mathcal{V}} = \|\frac{1}{\sqrt{a}} v\|_{L^2(\Omega)}$ for all $v \in H(\mathrm{div};\Omega)$. We seek $u \in L^2(\Omega)/\mathbf{R}$ and $\sigma \in \{v \in H(\mathrm{div};\Omega)\colon n \cdot v = 0, \text{ on } \partial\Omega\}$ such that

$$\left(\frac{1}{a}\sigma, v\right) + (u, \nabla \cdot v) = 0, \tag{4.1}$$

$$(-\nabla \cdot \sigma, w) = (f, w) \tag{4.2}$$

for all $w \in L^2(\Omega)/\mathbf{R}$ and $v \in \{v \in H(\mathrm{div};\Omega)\colon n \cdot v = 0, \text{ on } \partial\Omega\}$. Here $a \in L^\infty(\Omega)$ is positive bounded away from zero and $f \in L^2(\Omega)$ such that $\int_\Omega f \mathrm{d}x = 0$. Existence and uniqueness are guaranteed by the inf-sup condition; see, e.g., Theorem 2.34 in [8].

4.1. Notations. We let $\mathcal{W}_c = \mathcal{P}_H$ and $\mathcal{V}_c = \mathcal{RT}_H$, where $\mathcal{P}_H$ is the space of piecewise constants and $\mathcal{RT}_H = \mathrm{span}(\{\boldsymbol{\phi}_i\}_{i\in\mathcal{N}})$ is the space of lowest order Raviart–Thomas finite elements on the coarse mesh $\mathcal{K}_H$; see [6, [17], 20]. The Raviart–Thomas basis functions are defined over faces $\{F_i\}_{i\in\mathcal{N}}$ with normals $\{n_i\}_{i\in\mathcal{N}}$. In particular $F_i = \{x\colon n_i \cdot \boldsymbol{\phi}_i = 1\}$ for all $i \in \mathcal{N}$. We further let $\mathcal{P}_h(\omega_i)$ be the space of piecewise constants and $\mathcal{RT}_h(\omega_i)$ be the space of lowest order Raviart–Thomas basis functions on the fine subgrids $\mathcal{K}_h(\omega_i)$.

Let $P_H\colon L^2(\Omega) \to \mathcal{P}_H$ be the L^2 projection onto piecewise constants on the coarse mesh $\mathcal{K}_H$, and let $P_{h,\omega}\colon L^2(\omega) \to \mathcal{P}_h(\omega)$ be the corresponding projection onto $\mathcal{P}_h(\omega)$. We recall that $\|v - P_H v\|_{L^2(K)} \leq C H_K \|\nabla v\|_{L^2(K)}$ for all $K \in \mathcal{K}_H$ and that $\|v - P_{h,\omega} v\|_{L^2(K)} \leq C h_K \|\nabla v\|_{L^2(K)}$ for all $K \in \mathcal{K}_h(\omega)$. We let $\Pi_{\mathrm{RT}}\colon H(\mathrm{div};\Omega) \to \mathcal{RT}_H$ be the Raviart–Thomas interpolant onto the coarse space $\mathcal{RT}_H$; see, e.g., [8]. We will use the hierarchical split between the pairs $\mathcal{V}_c$,$\mathcal{W}_c$ and $\mathcal{V}_f$, $\mathcal{W}_f$. We let $\mathcal{P}_h^f(\omega_i) = \{v \in \mathcal{P}_h(\omega_i)\colon P_H v = 0\}$ and $\mathcal{RT}_h^f(\omega_i) = \{v \in \mathcal{RT}_h(\omega_i)\colon \Pi_{\mathrm{RT}} v = 0 \text{ and } n \cdot v = 0 \text{ on } \partial\omega_i\}$. Using the inclusion operators P_H and Π_{RT}, we also define the total fine scale spaces, $\mathcal{W}_f = \{w \in \mathcal{W}\colon P_H w = 0\}$ and $\mathcal{V}_f = \{v \in \mathcal{V}\colon \Pi_{\mathrm{RT}} v = 0\}$.

4.2. Coarse and fine scale equations. Since the solution has two components, we include all details when deriving the method, starting from the abstract framework. Again we first study the different kinds of local problems we need to solve, i.e., (2.9), (2.10), and (2.12). We immediately note that we have a symmetric problem; i.e., $\mathcal{T}^* = \mathcal{T}$. We also note that the local solutions will have two components, just like the original problem. We denote these by $\{\mathcal{T}_u, \mathcal{T}_\sigma\}$ and $\{u_f, \sigma_f\}$. We first translate (2.9) to the mixed setting and we get

$$\left(\frac{1}{a}\mathcal{T}_\sigma(w_c, v_c), v_f\right) + (\mathcal{T}_u(w_c, v_c), \nabla \cdot v_f) = -\left(\frac{1}{a}v_c, v_f\right) - (w_c, \nabla \cdot v_f), \tag{4.3}$$

$$(-\nabla \cdot \mathcal{T}_\sigma(w_c, v_c), w_f) = (\nabla \cdot v_c, w_f) \tag{4.4}$$

for all $v_c \in \mathcal{V}_c$, $w_c \in \mathcal{W}_c$, $v_f \in \mathcal{V}_f$, and $w_f \in \mathcal{W}_f$. The following lemma simplifies these equations.

LEMMA 4.1. *Using the hierarchical split, it holds that*

$$(w_c, \nabla \cdot v_f) = 0 \quad \textit{for all } w_c \in \mathcal{P}_H, \quad v_f \in \mathcal{V}_f, \tag{4.5}$$

$$(\nabla \cdot v_c, w_f) = 0 \quad \textit{for all } v_c \in \mathcal{RT}_H, \quad w_f \in \mathcal{W}_f. \tag{4.6}$$

Proof. See section 8.4. ∎

Since $(w_c, \nabla \cdot v_f) = 0$ and $(\nabla \cdot v_c, w_f) = 0$, according to Lemma 4.1, the system (4.3) becomes independent of w_c—i.e., $\mathcal{T}_\sigma(w_c, v_c) = \mathcal{T}_\sigma v_c$—and $\mathcal{T}_u(w_c, v_c) = \mathcal{T}_u v_c$ solves

$$\left(\frac{1}{a}\mathcal{T}_\sigma v_c, v_f\right) + (\mathcal{T}_u v_c, \nabla \cdot v_f) = -\left(\frac{1}{a}v_c, v_f\right), \tag{4.7}$$

$$(-\nabla \cdot \mathcal{T}_\sigma v_c, w_f) = 0 \tag{4.8}$$

for all $v_c \in \mathcal{V}_c$, $v_f \in \mathcal{V}_f$, and $w_f \in \mathcal{W}_f$. The equation corresponding to (2.12) reads as follows: find $u_f \in \mathcal{W}_f$ and $\sigma_f \in \mathcal{V}_f$ such that

$$\left(\frac{1}{a}\sigma_f, v_f\right) + (u_f, \nabla \cdot v_f) = 0, \tag{4.9}$$

$$(-\nabla \cdot \sigma_f, w_f) = (f, w_f) \tag{4.10}$$

for all $v_c \in \mathcal{V}_c$ and $v_f \in \mathcal{V}_f$. We derive the coarse and fine scale equations in the following lemma.

LEMMA 4.2. *Let $u = u_c + \mathcal{T}_u\sigma_c + u_f$ and $\sigma = \sigma_c + \mathcal{T}_\sigma\sigma_c + \sigma_f$ be unique solutions to* (4.1) *and* (4.2). *Then $\{u, \sigma\}$ is also unique solution to the following equivalent problem: find $\{u, \sigma\}$ such that*

$$\left(\frac{1}{a}(\sigma_c + \mathcal{T}_\sigma\sigma_c + \sigma_f), v_c + \mathcal{T}_\sigma v_c\right) + (u_c, \nabla \cdot v_c) - (\nabla \cdot \sigma_c, w_c) = (f, w_c), \tag{4.11}$$

$$\left(\frac{1}{a}\sigma_f, v_f\right) + (u_f, \nabla \cdot v_f) - (\nabla \cdot \sigma_f, w_f) = (f, w_f) \tag{4.12}$$

for all $v_c \in \mathcal{V}_c$, $w_c \in \mathcal{W}_c$, $v_f \in \mathcal{V}_f$, and $w_f \in \mathcal{W}_f$.

Proof. See section 8.5. ∎

4.3. Discrete method. We are ready to present the discrete method. We let $\tilde{\mathcal{T}}_\sigma v_c = \sum_{i\in\mathcal{N}} v_c^i \tilde{\mathcal{T}}_\sigma \boldsymbol{\phi}_i$, for all $v_c \in \mathcal{V}_c$, and $\tilde{\mathcal{T}}_u v_c = \sum_{i\in\mathcal{N}} v_c^i \tilde{\mathcal{T}}_u \boldsymbol{\phi}_i$. We formulate the localized fine scale problems in the following way: find $\tilde{\mathcal{T}}_u\boldsymbol{\phi}_i \in \mathcal{P}_h^f(\omega_i)$ and $\tilde{\mathcal{T}}_\sigma\boldsymbol{\phi}_i \in \mathcal{RT}_h^f(\omega_i)$ such that

$$\left(\frac{1}{a}\tilde{\mathcal{T}}_\sigma\boldsymbol{\phi}_i, v_f\right) + (\tilde{\mathcal{T}}_u\boldsymbol{\phi}_i, \nabla \cdot v_f) = -\left(\frac{1}{a}\boldsymbol{\phi}_i, v_f\right) \qquad \text{for all } v_f \in \mathcal{RT}_h^f(\omega_i), \tag{4.13}$$

$$(-\nabla \cdot \tilde{\mathcal{T}}_\sigma \boldsymbol{\phi}_i, w_f) = 0 \qquad \text{for all } w_f \in \mathcal{P}_h^f(\omega_i). \tag{4.14}$$

Here we have used a different technique to localize the right-hand side compared to the convection-diffusion-reaction equation. Instead of introducing a partition of unity, we use the split $\sigma_c = \sum_{i\in\mathcal{N}} \sigma_c^i \boldsymbol{\phi}_i$. This choice is very natural and easy to compute, but a similar approach for the convection-diffusion-reaction equation would lead to problems bounding $\tilde{\mathcal{T}}$ in Lemma 3.2. This problem does not appear in the mixed formulation. It would be possible to introduce a partition of unity for the mixed method as well, but it would lead to a method that is more complicated to implement. For the second fine scale equation, we still need a partition of unity in the scalar variable. This can be done in many ways, but we choose $\{\psi_i\}_{i\in\mathcal{N}}$ so that $\psi_i = \frac{1}{m}\text{supp}(\boldsymbol{\phi}_i)$, where m is the number of faces on each element K. This is a partition on unity since m members of the partition will be nonzero on every coarse element and therefore it sums up to one everywhere. Furthermore $\psi_i \in \mathcal{W}_c$ and $\psi_i w_c \in \mathcal{W}_c$ for all $w_c \in \mathcal{W}_c$. We get

$$\left(\frac{1}{a}\boldsymbol{\Sigma}_{f,i}, v_f\right) + (U_{f,i}, \nabla \cdot v_f) = 0 \qquad \text{for all } v_f \in \mathcal{RT}_h^f(\omega_i), \tag{4.15}$$

$$(-\nabla \cdot \boldsymbol{\Sigma}_{f,i}, w_f) = (f\psi_i, w_f) \qquad \text{for all } w_f \in \mathcal{P}_h^f(\omega_i). \tag{4.16}$$

Equations (4.13)–(4.14) and (4.15)–(4.16) are solvable since the data is known. The corresponding coarse scale equation reads as follows: find $U_c \in \mathcal{W}_c$ and $\boldsymbol{\Sigma}_c \in \mathcal{V}_c$ such that

$$\left(\frac{1}{a}(\boldsymbol{\Sigma}_c + \tilde{\mathcal{T}}_\sigma \boldsymbol{\Sigma}_c + \boldsymbol{\Sigma}_f), v_c + \tilde{\mathcal{T}}_\sigma v_c\right) + (U_c, \nabla \cdot v_c) + (\nabla \cdot \boldsymbol{\Sigma}_c, w_c) = -(f, w_c) \tag{4.17}$$

for all $v_c \in \mathcal{RT}_H$ and $w_c \in \mathcal{P}_H$. Note that this is the systematic technique to derive the symmetric version of the AVMS for mixed problems that was discussed briefly in Remark 2.1 in [15].

4.4. Error estimate. When deriving a posteriori error bounds for mixed methods using richer space for the flux than for the scalar variable, it is necessary to postprocess the solution to gain optimal order bounds; see [13], [21]. We note that $\boldsymbol{\Sigma}_c^i \tilde{\mathcal{T}}_u \boldsymbol{\phi}_i + U_{f,i}$, where $\boldsymbol{\Sigma}_c = \sum_{i\in\mathcal{N}} \boldsymbol{\Sigma}_c^i \boldsymbol{\phi}_i$ can be postprocessed using an approximation in the space of piecewise linear discontinuous functions. The postprocessed version of a function will be denoted with a star superscript *. In order to simplify notations, we let

$$U_{f,i}^{\text{tot}} = \boldsymbol{\Sigma}_c^i \tilde{\mathcal{T}}_u \boldsymbol{\phi}_i + U_{f,i}. \tag{4.18}$$

DEFINITION 4.1. *We let $\mathcal{P}_h^1(K)$ be the space of linear functions on $K \in \mathcal{K}_h(\omega_i)$ and define $U_{f,i}^{\text{tot},*}$ as follows:*

$$(U_{f,i}^{\text{tot},*}, v) = (U_{f,i}^{\text{tot}}, v) \qquad \textit{for all } v \in \mathcal{P}_h^1(K), \tag{4.19}$$

$$\begin{aligned} &(a\nabla U_{f,i}^{\text{tot},*}, \nabla w) = (\boldsymbol{\Sigma}_c^i(\boldsymbol{\phi}_i + \tilde{\mathcal{T}}_\sigma \boldsymbol{\phi}_i) + \boldsymbol{\Sigma}_{f,i}, \nabla w) \\ &\textit{for all } w \in \{w \in \mathcal{P}_h^1(K) : P_{h,K} w = 0\} \end{aligned} \tag{4.20}$$

for all fine elements $K \in \mathcal{K}_h(\omega_i)$.

We denote the postprocessed version $U^* = U_c + \sum_{i\in\mathcal{N}} U_{f,i}^{\mathrm{tot},*}$. Only the fine scale part of the solution is postprocessed. For technical reasons we need to introduce a special notation for a jump in a coarse function restricted to a coarse face; let

$$[U_{c,i}] = \begin{cases} [U_c] & \text{for all } x \in F_i, \\ 0 & \text{otherwise.} \end{cases} \tag{4.21}$$

Remember that $F_i = \{x : n_i^T \boldsymbol{\phi}_i = 1\}$, i.e., the coarse edge over which $\boldsymbol{\phi}_i$ is defined. We are now ready to derive an a posteriori error bound for this approximation.

THEOREM 4.1. *Let $u \in H^1(\Omega)/\mathbf{R}$ solve (4.1) and let $U^* = U_c + \sum_{i\in\mathcal{N}} U_{f,i}^{\mathrm{tot},*}$ be defined by (4.19) and (4.20). Then we have*

$$\begin{aligned} \left\| \frac{1}{\sqrt{a}}(\sigma - \Sigma) \right\|^2_{L^2(\Omega)} \le\ & C \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} \left\| \nabla U_{f,i}^{\mathrm{tot},*} - \frac{1}{a}(\Sigma_c^i(\boldsymbol{\phi}_i + \tilde{\mathcal{T}}\boldsymbol{\phi}_i) + \Sigma_{f,i}) \right\|^2_{L^2(K)} \\ & + C \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^{-1} \| [U_{c,i} + U_{f,i}^{\mathrm{tot},*}] \|^2_{L^2(\partial K)} \\ & + C \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^2 \| \psi_i f + \nabla \cdot \Sigma_c^i(\boldsymbol{\phi}_i + \tilde{\mathcal{T}}_\sigma \boldsymbol{\phi}_i) + \nabla \cdot \Sigma_{f,i} \|^2_{L^2(K)}. \end{aligned} \tag{4.22}$$

Proof. See section 8.6 ∎

Remark 4.1. The bound presented in (4.22) does not explicitly depend on the size of the subdomains. The first sum will decrease as h decreases, since the postprocessed gradients will approximate the flux better and better. The third term does depend on h explicitly and the second (jump) term will also decrease with h; see [13], [21]. However, there is a hidden effect in this term that will measure the influence of the restriction to subdomains. On the boundary $\partial\omega_i$ we have $[U_{c,i} + U_{f,i}^{\mathrm{tot},*}] = U_{f,i}^{\mathrm{tot},*}$; i.e., we have nothing to subtract with on the other side and the coarse part vanishes due to its definition. This means that this term can only be decreased by letting the patch increase, which will make $U_{f,i}^{\mathrm{tot},*}$ smaller. Note that we have the freedom to add any constant to $U_{f,i}^{\mathrm{tot},*}$. We use this to minimize its value on the boundary $\partial\omega_i$ by subtracting the mean.

5. Adaptive algorithms. Given an a posteriori error bound, it is natural to construct an adaptive algorithm in order to automatically tune the critical method parameters, i.e., the fine scale mesh sizes and the patch sizes. The refinement does not need to be uniform since we can calculate indicators for individual fine elements on the patch. However, in this section we will present a simple algorithm that uses different but uniform meshes on the patches.

We have two types of error indicators. One is defined in the interior of the patches I_{ω_i}, which will primarily depend on the fine scale mesh size, and one is defined on the boundaries of the patches $I_{\partial\omega_i}$, which will primarily depend on the size of the patches. For the convection-diffusion-reaction model problem, we have

$$\begin{aligned} I_{\omega_i}^2 = & \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^2 \| \boldsymbol{\phi}_i(f - \mathcal{L}U_c) - \mathcal{L}(U_{f,i} + \tilde{\mathcal{T}}_i U_c) \|^2_{L^2(K)} \\ & + \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K \| [a\boldsymbol{\phi}_i \partial_n U_c + a\partial_n(U_{f,i} + \tilde{\mathcal{T}}_i U_c)] \|^2_{L^2(\partial K \setminus \partial\omega_i)}, \end{aligned} \tag{5.1}$$

$$I^2_{\partial\omega_i} = \sum_{i\in\mathcal{N}} \max_{K\in\mathcal{K}_H(\omega_i)} H_K \|\Sigma^i_f(U)\|^2_{L^2(\partial\omega_i\setminus\partial\Omega)}, \tag{5.2}$$

where $\Sigma^i_f(U)$ is defined in the statement of Theorem 3.1. Here we have dropped the unknown constants. For the mixed model problem we have

$$\begin{aligned} I^2_{\omega_i} = &\sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} \left\| \nabla U^{\mathrm{tot},*}_{f,i} - \frac{1}{a}\left(\Sigma^i_c(\boldsymbol{\phi}_i + \tilde{\mathcal{T}}\boldsymbol{\phi}_i) + \Sigma_{f,i}\right)\right\|^2_{L^2(K)} \\ &+ \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^{-1} \|[U_{c,i} + U^{\mathrm{tot},*}_{f,i}]\|^2_{L^2(\partial K\setminus\partial\omega_i)} \\ &+ \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^2 \|\psi_i f + \nabla\cdot\Sigma^i_c(\boldsymbol{\phi}_i + \tilde{\mathcal{T}}_\sigma\boldsymbol{\phi}_i) + \nabla\cdot\Sigma_{f,i}\|^2_{L^2(K)}, \end{aligned} \tag{5.3}$$

$$I^2_{\partial\omega_i} = \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^{-1} \| U^{\mathrm{tot},*}_{f,i}\|^2_{L^2(\partial K\cap\partial\omega_i)}. \tag{5.4}$$

These error indicators are all easy and cheap to compute.

We treat the contributions from the boundary of the patches separately since this error mainly depends on the size of the subdomain ω_i. We are interested in creating an adaptive algorithm that automatically improves the solution in an iterative fashion based on an error estimate that scales correctly in the parameters of interest. The main goal is therefore not to calculate a good approximation of the error. This means that we are more interested in how the two indicators compare to each other than of their absolute value. This motivates us to drop the constants in the error estimates when constructing the indicators. We present the adaptive strategy in Algorithm 1. The number $|\mathcal{N}|$ is the number of coarse edges in the mesh (and therefore the number of local

Algorithm 1. Adaptive algorithm.

```
Construct initial coarse mesh with mesh size H
Let h^{ω_i} = H/2 and L(ω_i) = 1
while Σ_{i∈N}(I_{ω_i} + I_{∂ω_i}) > TOL do
  Calculate the solution to the primal problem, using the multiscale method, given {h^{ω_i}}_{i∈N} and {L(ω_i)}_{i∈N}
  Calculate the error indicators for each patch, {I_{ω_i}}_{i∈N} and {I_{∂ω_i}}_{i∈N}
  for i ∈ N do
    if I_{ω_i} > TOL/(2|N|), then
      h^{ω_i} := h^{ω_i}/2
    end if
    if I_{∂ω_i} > TOL/(2|N|), then
      L(ω_i) := L(ω_i) + 1
    end if
  end for
end while
```

problems that needs to be solved), h^{ω_i} and $L(\omega_i)$ represent local mesh resolution and number of layers, and TOL is a prescribed tolerance that the error indicators should meet. The stopping criteria can be improved by adding a restriction on the number of unknowns in each subproblem. The tolerance can also be replaced by a condition depending on the computational cost, if this is preferred.

6. Numerical examples. For the numerical examples we consider the Poisson equation on mixed form: find the pressure $u \in L^2(\Omega)/\mathbf{R}$ and flux $\sigma \in \{v \in H(\mathrm{div};\Omega)\colon n\cdot v = 0 \text{ on } \partial\Omega\}$ such that

$$\left(\frac{1}{a}\sigma, v\right) + (u, \nabla \cdot v) = 0, \tag{6.1}$$

$$(-\nabla \cdot \sigma, w) = (f, w) \tag{6.2}$$

for all $w \in L^2(\Omega)/\mathbf{R}$ and $v \in \{v \in H(\mathrm{div}; \Omega) : n \cdot v = 0, \text{ on } \partial\Omega\}$. Initially we consider three different types of permeability a, illustrated in Figure 6.1. The permeability is either constant, random (between 0 and 1 positive) on the fine scale, or taken for the tenth Society of Petroleum Engineering (SPE) comparative solution project modeling an oil reservoir. In the end of the section we focus on the oil reservoir application and consider two test cases from the SPE data set.

6.1. Solutions to local problems. We consider (4.13) and (4.14) with $\Omega = [0,1] \times [0,1]$ and $\mathcal{K}_H$ consisting of 8×8 square elements. We let $\omega_i^3 = [0, 3/4] \times [1/8, 3/4]$ be the three layer patch centered around the coarse face $\{(x, y) : x = 3/8, y \in (3/8, 4/8)\}$. We let ϕ_i be the coarse Raviart–Thomas basis function associated with the coarse edge and plot $|\tilde{\mathcal{T}}\phi_i|$, i.e., the absolute value of the flux part of the solution to local problem i. In Figure 6.2 we plot of the permeability in the region

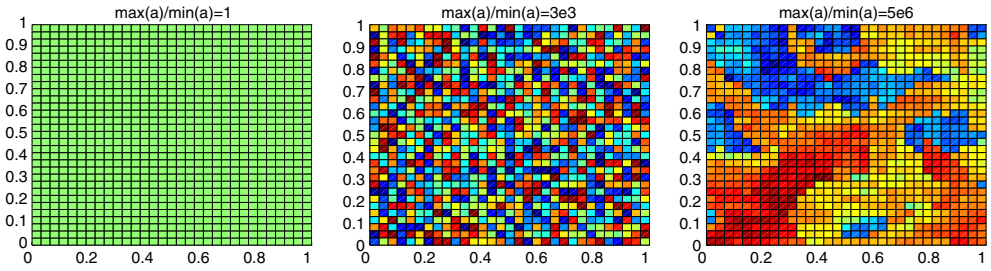

FIG. 6.1. *Constant permeability (left), a sample of a uniform distribution between* 0 *and* 1 *on the fine scale (middle), part of the tenth SPE comparative solution project data set (right).*

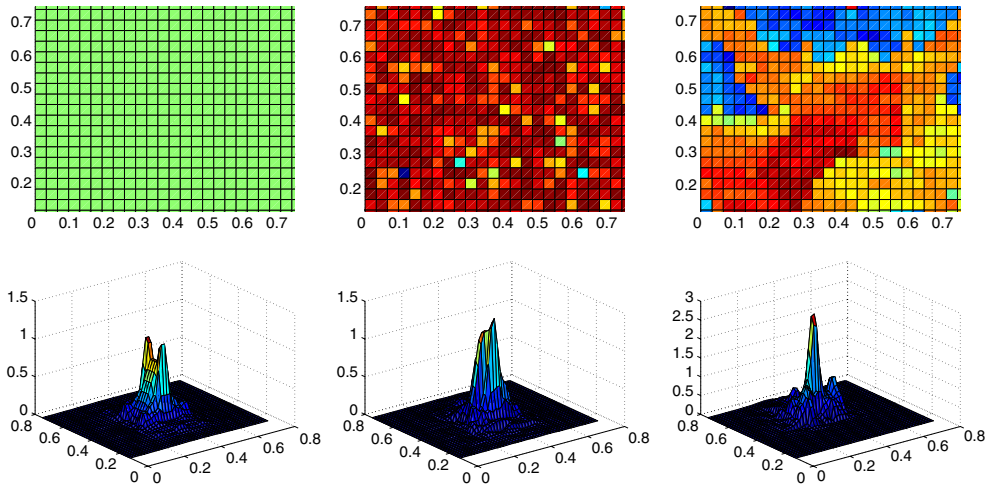

FIG. 6.2. *Above we plot the local permeability on a three layer patch surrounding the edge between* $(3/8, 3/8)$ *and* $(3/8, 4/8)$ *for, constant (left), one realization of a uniform distributed between* 0 *and* 1 *(middle), and part of the SPE data set (right). Below we plot the absolute value of the flux part of the corresponding local solutions,* $|\tilde{\mathcal{T}}_\sigma \phi_i|$.

where the computation is performed and (below) we plot the local solutions. We note the quick decay of the basis functions for all three types of permeability.

6.2. Convergence of global solution. We let $f = -1$ in the lower left corner $(0 \leq x, y \leq 1/32)$ and $f = 1$ in the upper right corner $(31/32 \leq x, y \leq 1)$; otherwise $f = 0$ and a are given by the three examples in Figure 6.1. Furthermore we let the reference mesh have 32×32 rectangular elements and the coarse mesh have 8×8 rectangular elements. We solve local problems with the same number of layers (L) and same resolution $h = H/4$ for all coarse Raviart–Thomas basis functions and compare the result with the reference solution in Figure 6.3. We get exponential decay in the relative error in energy norm $\|\frac{1}{\sqrt{a}}(\sigma_{\text{ref}} - \Sigma)\|_{L^2(\Omega)} / \|\frac{1}{\sqrt{a}}\sigma_{\text{ref}}\|_{L^2(\Omega)}$ in all three cases, where σ_{ref} is the fine scale reference solution. One layer seems to be enough to get a good approximation for the first two (constant and random) examples, and two layers is enough for the third (SPE) example.

6.3. A simple adaptivity example. We keep the same right-hand side data f as in the previous example. We now let a be periodic with strong multiscale features in the left half of the computational domain and constant in the right half. In Figure 6.4 we plot a (left), which varies between 1 and 0.02 in the left half of the domain and is equal to 1 in the right half, and the pressure u (right). We apply Algorithm 1 with 8×8 elements in the coarse mesh. We perform three iterations of the algorithm and mark 30% entries both for refinement and increased number of layers in each iteration. We plot the final pattern of refinements (left) and layers (right) in Figure 6.5. For each coarse element we take the mean value of the number of refinements and layers from the four local problems corresponding to the four edges. One sees that the most computational work is put in the left side of the computational domain and to regions where f is nonzero. We also plot the relative error in energy norm, compared to a reference solution computed using 64×64 elements, after each iteration in Figure 6.6. The relative error for standard

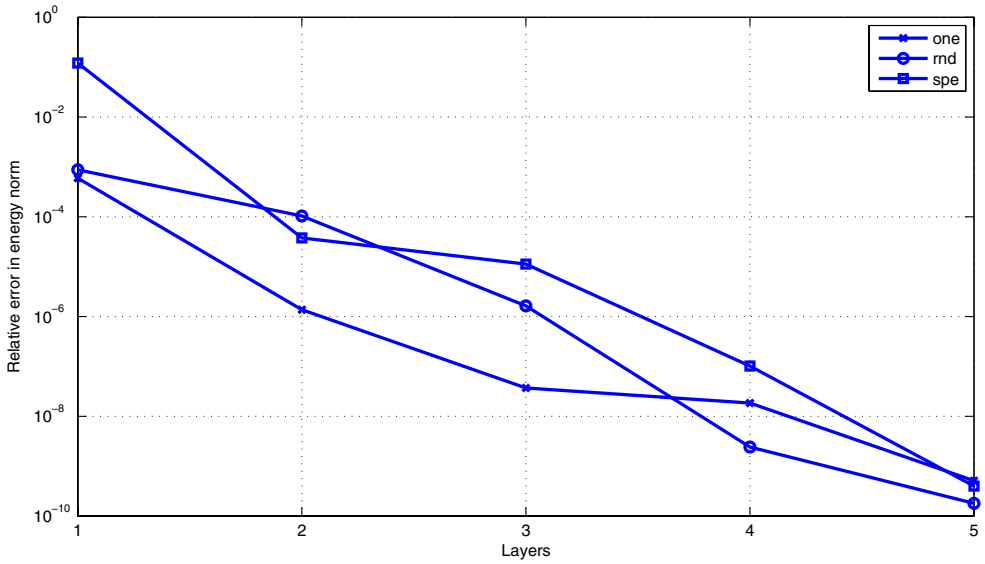

FIG. 6.3. *We plot the relative errors in energy norm* $\|\frac{1}{\sqrt{a}}(\sigma_{\text{ref}} - \Sigma)\|_{L^2(\Omega)} / \|\frac{1}{\sqrt{a}}\sigma_{\text{ref}}\|_{L^2(\Omega)}$ *compared to reference solutions* σ_{ref} *for the three cases, constant (one), random (rnd), and SPE (spe) permeability. We note that we get exponential decay in the error for all three cases.*

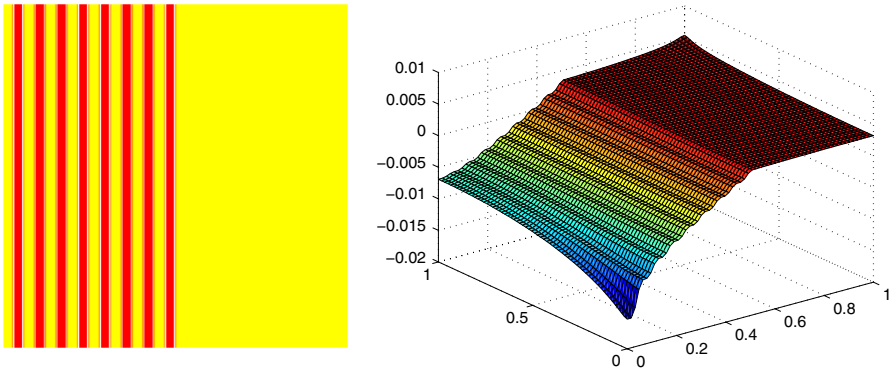

FIG. 6.4. *We plot the diffusion a (left), which varies between* 1 *and* 0.02*, and the pressure u (right)*

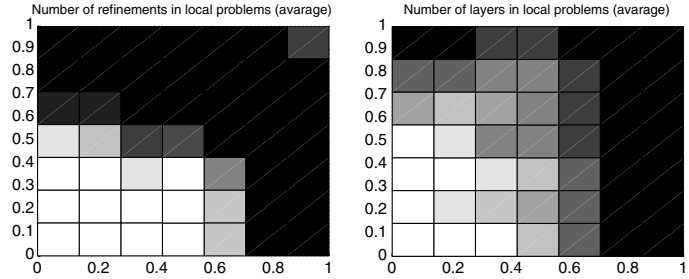

FIG. 6.5. *For each coarse element the mean of the number of local refinements (left) and layers (right) for the local problems associated with the four edges of the element is plotted. Black corresponds to one refinement or one layer and white corresponds to three refinements or layers. Since we take averages over all four corresponding edges, for each coarse element, we get fractional number between one and three, indicating the level of accuracy.*

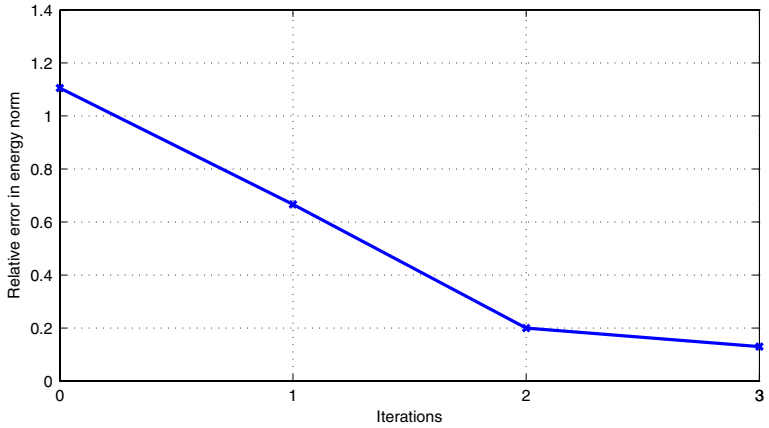

FIG. 6.6. *Relative error in energy norm* $\|\frac{1}{\sqrt{a}}\cdot\|_{L^2(\Omega)}$ *compared to reference solution after each iteration in the adaptive algorithm. Zero iterations represents the standard Galerkin solution on the coarse mesh.*

Galerkin (0 layers) is over 100% because the fine scale features are not resolved. After three iterations the error is down to 13%.

6.4. Application to oil reservoir simulation. We will show numerical examples in two spatial dimensions using permeability from the tenth SPE comparative solution project. A frequently used method in oil recovery is to inject water into the reservoir in order to move the oil to a producing well. Quick numerical simulation is important in this process in order to make correct decisions. The two phase flow in the oil reservoir can be modeled by an elliptic pressure equation coupled to a hyperbolic saturation equation

$$\nabla \cdot (\lambda(S)\mathbb{K}\nabla p) = h, \tag{6.3}$$

$$\frac{\partial S}{\partial t} + \mathbf{v} \cdot \nabla f(S) = h_w, \tag{6.4}$$

where $\mathbb{K}$ is the permeability tensor, S is the water saturation, p is the pressure, λ is the total mobility, f is the fractional flow of water, $\mathbf{v}$ is the total velocity, and h and h_w are source/sink terms. See, e.g., [3] for a more extensive discussion on two phase flow model problems. We consider the pressure equation (6.3) and note that if we let $\lambda(S) = 1$, $a = \mathbb{K}$, $u = p$, $\sigma = a\nabla u$, $h = -f$, and write the equations on mixed form, we exactly get our model problem (6.1)–(6.2).

In Figure 6.7 we plot the permeability a from layer 1 and 50 in the SPE data set (in log scale). In Figure 6.8 we plot the corresponding solutions for the pressure. The data

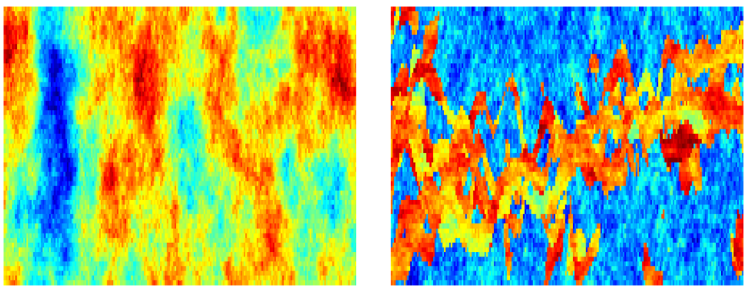

FIG. 6.7. *Permeability from layer* 1 *and* 50 *of the tenth SPE comparative solution project in log scale.*

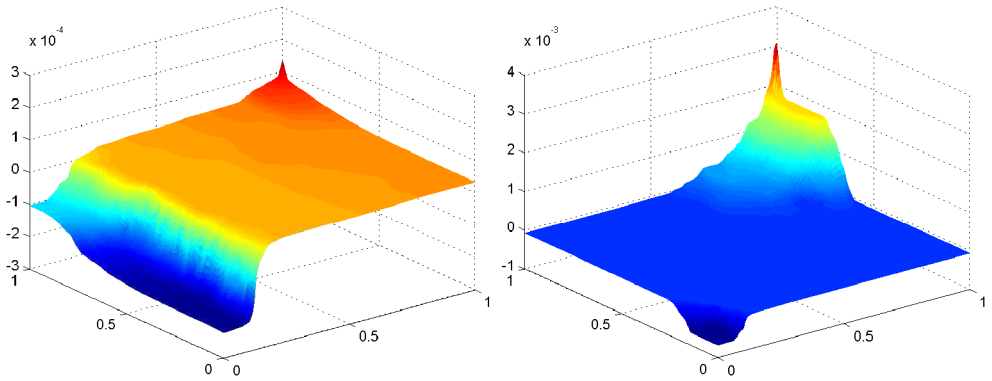

FIG. 6.8. *Reference solution (pressure) for layer* 1 *and* 50 *in the SPE data set.*

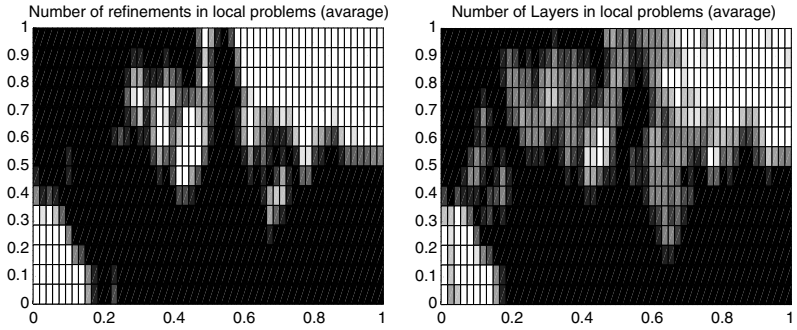

FIG. 6.9. *Refinements and layers as in Figure* 6.5.

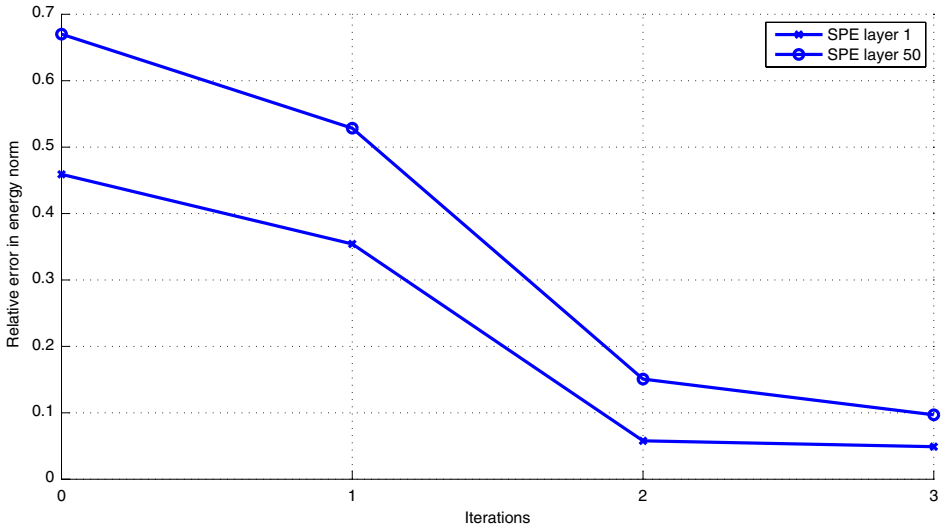

FIG. 6.10. *Relative error in energy norm* $\|\frac{1}{\sqrt{a}}\cdot\|_{L^2(\Omega)}$ *compared to reference solution after each iteration in the adaptive algorithm. Zero iterations represents the standard Galerkin solution on the coarse mesh.*

was chosen to be $f = -1$ in the lower left corner $(0 \le x \le 1/220, 0 \le y \le 1/60)$, $f = 1$ in the upper right corner $(219/220 \le x \le 1, 59/60 \le y \le 1)$, and $f = 0$ otherwise. Next we apply the adaptive algorithm, Algorithm 1, to compute an approximate solution. We start with one layer patches and one refinement for all local problems. We use 55×15 coarse elements and a reference mesh with 440×120 elements, which we measure the error against. In each iteration we again mark 30% of the entries for refinement both in terms of layers and local mesh size. In Figure 6.9 we show the number of refinements and layers after three iterations in the adaptive algorithm. We see that most effort is put in areas close to the injector and the producer. The relative error in energy norm after each iteration is plotted in Figure 6.10. We note that the adaptive algorithm manages to reduce the error considerably and that a majority of the patches still use only one layer and one refinement after three iterations.

7. Conclusions. In this paper we present a systematic technique for constructing multiscale methods for a wide range of partial differential equations, including diffusion

and convection dominated problems. We show by examples that the methods derived give accurate solutions for different data types. We also provide a posteriori error estimates for two distinct model problems and present an adaptive algorithm for automatic tuning of the method parameters. In the numerical section we apply the method to a model problem frequently used in oil reservoir simulation with promising results. We conclude that the method appears to be robust and gives promising results for a range of different multiscale data. The adaptive nature of the method makes it possible to avoid solving expensive local problems everywhere in the computational domain. Instead, different resolution can be used in different regions, and the adaptive algorithm governs these delicate choices automatically. This is the main strength of this method compared to competing methods.

8. Proofs of theoretical results. In this section we collect proofs of the theoretical results of the paper.

8.1. Lemma 3.1.

Proof. In three spatial dimensions,

$$\|\boldsymbol{\phi}_i v_f\|^2_{H^1_0(\Omega)} = (a\boldsymbol{\phi}_i \nabla v_f, \nabla(\boldsymbol{\phi}_i v_f)) + (a v_f \nabla \boldsymbol{\phi}_i, \nabla(\boldsymbol{\phi}_i v_f)) \tag{8.1}$$

$$\leq \|\boldsymbol{\phi}_i\|_{L^\infty(\Omega)} \|v_f\|_{H^1_0(\Omega)} \|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)}$$

$$+ C\|\boldsymbol{\phi}_i\|_{W^{1,3}(\mathrm{supp}(\boldsymbol{\phi}_i))} \|v_f\|_{L^6(\mathrm{supp}(\boldsymbol{\phi}_i))} \|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)} \tag{8.2}$$

$$\leq \|v_f\|_{H^1_0(\Omega)} \|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)} + C\|\boldsymbol{\phi}_i\|_{W^{1,3}(\mathrm{supp}(\boldsymbol{\phi}_i))} \|v_f\|_{H^1(\mathrm{supp}(\boldsymbol{\phi}_i))} \|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)} \tag{8.3}$$

$$\leq \|v_f\|_{H^1_0(\Omega)} \|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)} + C\left(\frac{|\mathrm{supp}(\boldsymbol{\phi}_i)|}{\min\limits_{K_H \in \mathrm{supp}(\boldsymbol{\phi}_i)} H_K^3}\right)^{1/3} \|v_f\|_{H^1_0(\Omega)} \|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)} \tag{8.4}$$

$$\leq C\|v_f\|_{H^1_0(\Omega)} \|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)}; \tag{8.5}$$

i.e., $\|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)} \leq C\|v_f\|_{H^1_0(\Omega)}$ for all $v_f \in \mathcal{V}_f$. In two spatial dimensions a similar argument holds, but here $\|\boldsymbol{\phi}_i v_f\|_{H^1_0(\Omega)} \leq CH^{-\epsilon}\|v_f\|_{H^1_0(\Omega)}$ for any $\epsilon > 0$. ∎

8.2. Lemma 3.2.

Proof. We study $\tilde{\mathcal{T}}_i^* v_c \in \mathcal{V}_f^h(\omega_i)$ and use the assumption $\alpha^{-1}\|v\|^2_{H^1_0(\Omega)} \leq a(v,v)$ and (3.9) to get

$$\alpha^{-1}\|\tilde{\mathcal{T}}_i^* v_c\|^2_{H^1(\omega_i)} \leq a(\tilde{\mathcal{T}}_i^* v_c, \tilde{\mathcal{T}}_i^* v_c) = -a(\boldsymbol{\phi}_i \tilde{\mathcal{T}}_i^* v_c, v_c) \leq C\|v_c\|_{H^1(\mathrm{supp}(\boldsymbol{\phi}_i))} \|\boldsymbol{\phi}_i \tilde{\mathcal{T}}_i^* v_c\|_{H^1(\omega_i)} \tag{8.6}$$

$$\leq C\|v_c\|_{H^1(\mathrm{supp}(\boldsymbol{\phi}_i))} \|\tilde{\mathcal{T}}_i^* v_c\|_{H^1(\omega_i)}, \tag{8.7}$$

using Lemma 3.1. We conclude that $\|\tilde{\mathcal{T}}_i^* v_c\|_{H^1(\omega_i)} \leq C\|v_c\|_{H^1(\mathrm{supp}(\boldsymbol{\phi}_i))}$.; i.e.,

$$\|\tilde{\mathcal{T}}^* v_c\|^2_{H^1(\Omega)} \leq \sum_{i \in \mathcal{N}} \|\tilde{\mathcal{T}}_i^* v_c\|^2_{H^1(\omega_i)} \leq C \sum_{i \in \mathcal{N}} \|v_c\|^2_{H^1(\mathrm{supp}(\boldsymbol{\phi}_i))} \leq C\|v_c\|^2_{H^1(\Omega)}. \tag{8.8}$$

Note that the support of $\boldsymbol{\phi}_i$ overlaps, but the size is controlled by the number of elements that surrounds one node. ∎

8.3. Theorem 3.1.

Proof. We want to subtract interpolants on the fine meshes $\mathcal{K}_h(\omega_i)$ to get the optimal h dependency in the estimate. We will use the orthogonality of the method to get this. We start from the left-hand side and let $e = u - U$,

$$\alpha^{-1}\|e\|^2_{H^1(\Omega)} \le a(e,e) \tag{8.9}$$

$$= l(e) - a(U,e) \tag{8.10}$$

$$= l(e - \pi_c e - \tilde{\mathcal{T}}^*\pi_c e) - a(U, e - \pi_c e - \tilde{\mathcal{T}}^*\pi_c e) \tag{8.11}$$

$$= l((1-\pi_c)(e - \tilde{\mathcal{T}}^*\pi_c e)) - a(U,(1-\pi_c)(e - \tilde{\mathcal{T}}^*\pi_c e)) \tag{8.12}$$

$$\le \sup_{v\in H^1_0(\Omega)} \frac{l(v-\pi_c v) - a(U, v-\pi_c v)}{\|v\|_{H^1(\Omega)}} \|e - \tilde{\mathcal{T}}^*\pi_c e\|_{H^1(\Omega)}, \tag{8.13}$$

using coercivity and (3.12). We apply Lemma 3.2 with $v_c = \pi_c e$ and note that π_c is stable in $H^1_0(\Omega)$ (see [5]) to get $\|e - \tilde{\mathcal{T}}^*\pi_c e\|_{H^1(\Omega)} \le C\|e\|_{H^1(\Omega)}$; i.e.,

$$\|e\|_{H^1(\Omega)} \le C \sup_{v\in H^1_0(\Omega)} \frac{l(v-\pi_c v) - a(U, v-\pi_c v)}{\|v\|_{H^1(\Omega)}}. \tag{8.14}$$

We fix v and use the fine scale (3.8) and (3.10) to get

$$\begin{aligned} l(v-\pi_c v) - a(U, v-\pi_c v) &= \sum_{i\in\mathcal{N}} (f\boldsymbol{\phi}_i, v-\pi_c v) - a(U_c, (v-\pi_c v)\boldsymbol{\phi}_i), \\ &\sum_{i\in\mathcal{N}} - a(U_{f,i} + \tilde{\mathcal{T}}_i U_c, (v-\pi_c v)) \end{aligned} \tag{8.15}$$

$$\begin{aligned} &= \sum_{i\in\mathcal{N}} (f\boldsymbol{\phi}_i, v-\pi_c v - \pi_f^i v) - a(U_c, (v-\pi_c v - \pi_f^i v)\boldsymbol{\phi}_i), \\ &\sum_{i\in\mathcal{N}} - a(U_{f,i} + \tilde{\mathcal{T}}_i U_c, v-\pi_c v - \pi_f^i v), \end{aligned} \tag{8.16}$$

where $\pi_f^i : H^1_0(\Omega) \to \mathcal{V}_f^h(\omega_i)$ is the Scott–Zhang interpolant onto $\mathcal{V}_f^h(\omega_i)$. It is inconvenient to work with π_f^i since $v - \pi_c v$ does not vanish on the boundaries $\partial\omega_i$. We let $\tilde{\mathcal{V}}^h(\omega_i)$ be the space of continuous piecewise linear functions on $\mathcal{K}_h(\omega)$ with no extra assumption at the boundary. We let $\tilde{\pi} : H^1_0(\Omega) \to \tilde{\mathcal{V}}^h(\omega_i)$ be the corresponding Scott–Zhang interpolant and then define $\pi_f = \tilde{\pi} - \pi_c$. We have control over π_f since we have interpolation bound for $v - \pi_c v - \pi_f v$ on each subdomain ω_i. We therefore replace π_f^i with π_f by adding and subtracting the same term

$$\begin{aligned} l(v-\pi_c v) - a(U, v-\pi_c v) &= \sum_{i\in\mathcal{N}} (f\boldsymbol{\phi}_i, v-\pi_c v - \pi_f v) - a(U_c, (v-\pi_c v - \pi_f v)\boldsymbol{\phi}_i) \\ &- \sum_{i\in\mathcal{N}} a(U_{f,i} + \tilde{\mathcal{T}}_i U_c, v-\pi_c v - \pi_f v) \\ &+ \sum_{i\in\mathcal{N}} (f\boldsymbol{\phi}_i, \pi_f v - \pi_f^i v) - a(U_c, (\pi_f v - \pi_f^i)\boldsymbol{\phi}_i) - a(U_{f,i}, \pi_f v - \pi_f^i v) \end{aligned} \tag{8.17}$$

$$
\begin{aligned}
\leq & \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K \|\boldsymbol{\phi}_i(f-\mathcal{L}U_c) - \mathcal{L}(U_{f,i}+\tilde{\mathcal{T}}_i U_c)\|_{L^2(K)} h_K^{-1} \|v-\pi_c v-\pi_f v\|_{L^2(K)} \\
& + \sum_{i\in\mathcal{N}} \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^{1/2} \|[a\boldsymbol{\phi}_i \partial_n U_c + a\partial_n(U_{f,i}+\tilde{\mathcal{T}}_i U_c)]\|_{L^2(\partial K)} h_K^{-1/2} \|v-\pi_c v-\pi_f v\|_{L^2(\partial K)} \\
& + \left| \sum_{i\in\mathcal{N}} \langle \Sigma_f^i(U), \pi_f(v-\pi_c v)\rangle_{\partial\omega_i} \right|,
\end{aligned}
\tag{8.18}
$$

where we use Green's formula on fine elements, to move the gradient in the diffusion term to the left slot in the scalar product (see, e.g., Theorem 8.1 in [4]), together with the Cauchy–Schwartz inequality and the definition of $\Sigma_f^i(U)$ from the statement of the theorem. We can apply Cauchy–Schwartz inequality for sums and use the interpolation result

$$
\sum_{K\in\mathcal{K}_h(\omega_i)} h_K^{-2} \|v-\pi_c v-\pi_f v\|_{L^2(K)}^2 \leq C\|v\|_{H^1(\omega_i)}^2,
\tag{8.19}
$$

and the trace inequality followed by an interpolation estimate

$$
\begin{aligned}
\sum_{K\in\mathcal{K}_h(\omega_i)} h_K^{-1} \|v-\pi_c v-\pi_f v\|_{L^2(\partial K)}^2 \leq & \; C \sum_{K\in\mathcal{K}_h(\omega_i)} h_K^{-2} \|v-\pi_c v-\pi_f v\|_{L^2(K)}^2 \\
& + C \sum_{K\in\mathcal{K}_h(\omega_i)} \|v-\pi_c v-\pi_f v\|_{H^1(K)}^2
\end{aligned}
\tag{8.20}
$$

$$
\leq C\|v\|_{H^1(\omega_i)}^2
\tag{8.21}
$$

to get

$$
\begin{aligned}
l(v-\pi_c v) - a(U, v-\pi_c v) \leq & \; C\left(\sum\nolimits_{i\in\mathcal{N}} \sum\nolimits_{K\in\mathcal{K}_h(\omega_i)} \rho_K^i(U)\right)^{1/2} \left(\sum\nolimits_{i\in\mathcal{N}} \|v\|_{H^1(\omega_i)}^2\right)^{1/2} \\
& + |\langle \Sigma_f^i(U), \pi_f(v-\pi_c v)\rangle_{\partial\omega_i}|
\end{aligned}
\tag{8.22}
$$

$$
\begin{aligned}
\leq & \; C\left(\sum\nolimits_{i\in\mathcal{N}} \sum\nolimits_{K\in\mathcal{K}_h(\omega_i)} \rho_K^i(U)\right)^{1/2} \|v\|_{H^1(\Omega)} \\
& + |\langle \Sigma_f^i(U), \pi_f(v-\pi_c v)\rangle_{\partial\omega_i}|,
\end{aligned}
\tag{8.23}
$$

using the definition in the formulation of the theorem and that ω_i are built up by a few layers of coarse elements. After dividing the first term with $\|v\|_{H^1(\Omega)}$ and taking supremum, we get

$$
\|e\|_{H^1(\Omega)} \leq C\left(\sum\nolimits_{i\in\mathcal{N}} \sum\nolimits_{K\in\mathcal{K}_h(\omega_i)} \rho_K^i(U)\right)^{1/2} + \sup_{v\in H_0^1(\Omega)} \frac{\sum_{i\in\mathcal{N}} |\langle \Sigma_f^i(U), \pi_f(v-\pi_c v)\rangle_{\partial\omega_i}|}{\|v\|_{H^1(\Omega)}}.
\tag{8.24}
$$

It remains to bound the last term of (8.24). We use the same trace inequality as above (8.20) on the coarse mesh and the stability of the Scott–Zhang interpolant to get

$$\sum_{i\in\mathcal{N}}|\langle\Sigma_f^i(U),\pi_f v\rangle_{\partial\omega_i}| \le C\sum_{i\in\mathcal{N}}\max_{K\in\mathcal{K}_H(\omega_i)} H_K^{1/2}\|\Sigma_f^i(U)\|_{\partial\omega_i}\|v\|_{H^1(\omega_i)}, \tag{8.25}$$

$$\le C\left(\sum\nolimits_{i\in\mathcal{N}}\max\nolimits_{K\in\mathcal{K}_H(\omega_i)}\ H_K\|\Sigma_f^i(U)\|^2_{\partial\omega_i}\right)^{1/2}\|v\|_{H^1(\Omega)}. \tag{8.26}$$

The theorem follows immediately by combining (8.24) and (8.26). ∎

8.4. Lemma 4.1.

Proof. We start with the first equality,

$$(w_c,\nabla\cdot v_f)=\sum_{K\in\mathcal{K}_H} w_c|_K\int_K\nabla\cdot v_f\mathrm{d}x=\sum_{K\in\mathcal{K}_H} w_c|_K\int_{\partial K} n\cdot v_f\mathrm{d}s=0, \tag{8.27}$$

since $n\cdot v_f$ has mean value zero over each coarse face due to the hierarchical split. We turn to the next equality,

$$(\nabla\cdot v_c,w_f)=\sum_{K\in\mathcal{K}_H}\nabla\cdot v_c|_K\int_K w_f\mathrm{d}x=0, \tag{8.28}$$

since w_f has average zero over each coarse element due to construction. ∎

8.5. Lemma 4.2.

Proof. Plug $u=u_c+\mathcal{T}_u\sigma_c+u_f$, $\sigma=\sigma_c+\mathcal{T}_\sigma\sigma_c+\sigma_f$, $w=w_c+\mathcal{T}_u v_c+w_f$, and $v=v_c+\mathcal{T}_\sigma v_c+v_f$ into (4.1) and (4.2). Several terms vanish due to Lemma 4.1 and the relations (4.7)–(4.10). We note that $(\frac{1}{a}(\sigma_c+\mathcal{T}_\sigma\sigma_c),v_f)+(u_c+\mathcal{T}_u\sigma_c,\nabla\cdot v_f)=0$ for all $v_f\in\mathcal{V}_f$, $(w,\nabla\cdot\mathcal{T}_\sigma v_c)=0$ for all $v_c\in\mathcal{V}_c$ and all $w\in\mathcal{W}$, and $(-\nabla\cdot\sigma_f,\mathcal{T}_u v_c)-(f,\mathcal{T}_u v_c)=0$ for all $v_c\in\mathcal{V}_c$. The lemma follows immediately. ∎

8.6. Theorem 4.1.

Proof. The proof is identical to the proof of Corollary 3.1 [15], which is based on Theorem 3.1 also in [15]. The difference in the two cases is that a symmetric formulation is used in this paper. Since no interpolant needs to be subtracted in the flux variable (in particular not the coarse scale interpolant), stability with respect to the $\tilde{\mathcal{T}}$ operator does not come in to the error analysis as it did in the convection-diffusion-reaction case and does not effect the estimate derived Theorem 3.1 in [15]. This observation makes the proof directly applicable. ∎

REFERENCES

[1] R. A. Adams, *Sobolev Spaces*, Pure Appl. Math. 65, Academic Press, New York, 1975.

[2] T. Arbogast and S. L. Bryant, *A two-scale numerical subgrid technique for waterflood simulations*, SPE J., 7 (2002), pp. 446–457.

[3] K. Aziz and A. Settari, *Petroleum Reservoir Simulation*, Applied Science Publishers, London, 1979.

[4] D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, 2007.

[5] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[6] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[7] J. Chu, Y. Efendiev, V. Ginting, and T. Y. Hou, *Flow based oversampling technique for multiscale finite element methods*, Adv. Water Resour., 31 (2008), pp. 599–608.

[8] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Springer-Verlag, New York, 2004.

[9] 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), pp. 169–189.
[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), pp. 387–401.
[11] 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 (1998), pp. 3–24.
[12] 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), pp. 2313–2324.
[13] M. G. Larson and A. Målqvist, *A posteriori error estimates in the energy norm for mixed finite element methods*, Numer. Math., 108 (2008), pp. 487–500.
[14] M. G. Larson and A. Målqvist, *Adaptive variational multiscale method of convection-diffusion problems*, Comm. Numer. Methods Engrg., 25 (2009), pp. 65–79.
[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 (2009), pp. 1017–1042.
[16] P. Morin, R. H. Nochetto, and K. G. Siebert, *Local problems on stars: A posteriori error estimators, convergence, and performance,*, Math. Comp., 72 (2003), pp. 1067–1097.
[17] J.-C. Nédélec, *A new family of mixed finite elements in* $\mathbf{R}^3$, Numer. Math., 50 (1986), pp. 57–81.
[18] J. Nolen, G. Papanicolaou, and O. Pironneau, *A framework for adaptive multiscale methods for elliptic problems*, Multiscale Model. Simul., 7 (2008), pp. 171–196.
[19] J. M. Nordbotten, *Adaptive variational multiscale methods for multi-phase flow in porous media*, Multiscale Model. Simul., 7 (2008), pp. 1455–1473.
[20] P. Raviart and J. Thomas, *A mixed finite element method for second order elliptic problems*, in Mathematical Aspects of the Finite Element Method, Lecture Notes in Math. 606, Springer-Verlag, New York, 1977, pp. 292–315.
[21] C. Lovadina and R. Stenberg, *Energy norm a posteriori error estimates for mixed finite element methods*, Math. Comp., 75 (2006), pp. 1659–1674.