Mixed Multiscale Methods for Heterogeneous Elliptic Problems

Todd Arbogast

Abstract We consider a second order elliptic problem written in mixed form, i.e., as a system of two first order equations. Such problems arise in many contexts, including flow in porous media. The coefficient in the elliptic problem (the permeability of the porous medium) is assumed to be spatially heterogeneous. The emphasis here is on accurate approximation of the solution with respect to the scale of variation in this coefficient. Homogenization and upscaling techniques alone are generally inadequate for this problem. As an alternative, multiscale numerical methods have been developed. They can be viewed in one of three equivalent frameworks: as a Galerkin or finite element method with nonpolynomial basis functions, as a variational multiscale method with standard finite elements, or as a domain decomposition method with restricted degrees of freedom on the interfaces. We treat each case, and discuss the advantages of the approach for devising effective local multiscale methods. Included is recent work on methods that incorporate information from homogenization theory and effective domain decomposition methods.

1 Elliptic Systems with a Heterogeneous Coefficient

We consider a second order elliptic problem, which we write in *mixed form,* i.e., as the following system of two first order equations:

$$\mathbf{u} = -a_\varepsilon \nabla p \quad \text{in } \Omega \subset \mathbb{R}^d \quad \text{(Darcy's law)}, \tag{1}$$

$$\nabla \cdot \mathbf{u} = f \quad \text{in } \Omega \quad \text{(Conservation)}, \tag{2}$$

$$\mathbf{u} \cdot \nu = 0 \quad \text{on } \partial\Omega. \tag{3}$$

Todd Arbogast
The University of Texas at Austin, Institute for Computational Engineering and Sciences, 1 University Station C0200, Austin, Texas 78712, USA; and The University of Texas at Austin, Mathematics Department, 1 University Station C1200, Austin, Texas 78712–0257, USA; e-mail: arbogast@ices.utexas.edu

The system models, e.g., incompressible, single phase flow in a porous medium in which case the unknowns are p, the fluid pressure, and $\mathbf{u}$, the Darcy velocity of the fluid, while the known parameters are a_ε, the medium permeability, and f, the source or sink term (i.e., the wells) [17, 18, 29]. The first equation is the empirical *Darcy Law* relating the velocity to the pressure gradient, the second equation expresses the mass conservation principle, and we take a boundary condition representing no normal flow for simplicity of exposition. In this discussion, we assume that a_ε is *heterogeneous* on a scale ε.

Our objective is to find an *accurate* approximation of $\mathbf{u}$ and p while we respect the principle of mass *conservation,* which is a critical property in many applications. In fact, we are most interested in an accurate, conservative approximation to $\mathbf{u}$, since the velocity controls the transport of mass, such as a contaminant in the groundwater or the macroscopic mixing of multiple phases.

We can rewrite the system (1)–(3) in *mixed variational form* as follows. Let $(\cdot,\cdot)_\omega$ denote the $L^2(\omega)$ or $(L^2(\omega))^d$ inner product, wherein we omit ω when it is Ω, and define the Hilbert spaces $H(\mathrm{div};\Omega) := \{\mathbf{v} \in (L^2(\Omega))^d : \nabla\cdot\mathbf{v} \in L^2(\Omega)\}$ and

$$\mathbf{V} := H_0(\mathrm{div};\Omega) := \{\mathbf{v} \in H(\mathrm{div};\Omega) : \mathbf{v}\cdot\nu = 0 \text{ on } \partial\Omega\},$$

where $\|\mathbf{v}\|_{\mathbf{V}}^2 = \|\mathbf{v}\|_0^2 + \|\nabla\cdot\mathbf{v}\|_0^2$, $\|\psi\|_0^2 = (\psi,\psi)$. Using integration by parts to rewrite

$$-(\nabla p,\mathbf{v}) = (p,\nabla\cdot\mathbf{v}),$$

the problem is equivalent to

Find $p \in W = L^2(\Omega)/\mathbb{R}$ and $\mathbf{u} \in \mathbf{V}$ such that

$$(a_\varepsilon^{-1}\mathbf{u},\mathbf{v}) - (p,\nabla\cdot\mathbf{v}) = 0 \qquad \forall\, \mathbf{v} \in \mathbf{V} \quad \text{(Darcy's law)}, \tag{4}$$

$$(\nabla\cdot\mathbf{u},w) = (f,w) \quad \forall\, w \in W \quad \text{(Conservation)}. \tag{5}$$

We remark that the mixed form preserves the conservation equation, and so allows locally conservative approximations.

We have a saddle-point problem, since it has both positive and negative eigenvalues. There is a well-developed abstract theory for the well-posedness of such mixed variational forms [13, 22, 26, 21].

Theorem 1. [Babuška, 1973; Brezzi, 1974] *For the abstract saddle-point problem*

Find $p \in W$ and $\mathbf{u} \in \mathbf{V}$ such that

$$A(\mathbf{u},\mathbf{v}) - (p,\nabla\cdot\mathbf{v}) = G(\mathbf{v}) \quad \forall\, \mathbf{v} \in \mathbf{V},$$
$$(w,\nabla\cdot\mathbf{u}) = F(w) \quad \forall\, w \in W,$$

suppose A is a continuous, symmetric bilinear form, coercive on $\mathbf{V}\cap\ker(\nabla\cdot)$, and that there exists $\gamma > 0$ such that

$$\inf_{w\in W}\sup_{\mathbf{v}\in\mathbf{V}} \frac{(w,\nabla\cdot\mathbf{v})}{\|w\|_W\,\|\mathbf{v}\|_{\mathbf{V}}} \ge \gamma. \tag{6}$$

Then there exists a unique solution $(p,\mathbf{u}) \in W \times \mathbf{V}$*, and*

$$\|p\|_W + \|\mathbf{u}\|_{\mathbf{V}} \leq C\{\|F\|_{W^*} + \|G\|_{\mathbf{V}^*}\}.$$

In our case, we satisfy the inf-sup condition (6), and

$$A(\mathbf{u},\mathbf{v}) = (a_\varepsilon^{-1}\mathbf{u},\mathbf{v})$$

is continuous and coercive provided that the tensor $a_\varepsilon \in (L^\infty(\Omega))^{d\times d}$ is uniformly positive definite: there are $\alpha_* > 0$ and $\alpha^* < \infty$ such that

$$\alpha_*|\xi|^2 \leq \xi^T a_\varepsilon(x)\xi \leq \alpha^*|\xi|^2 \quad \forall\, \xi \in \mathbb{R}^d,\ a.e.\ x \in \Omega. \tag{7}$$

Thus we have a well-posed problem for, say, $f \in L^2(\Omega)$.

The complication comes from the problem of scale. Because a_ε varies on the spatial scale ε,

$$|\mathbf{u}| = \mathscr{O}(1) \quad \text{but} \quad |D^k\mathbf{u}| = \mathscr{O}(\varepsilon^{-k}).$$

Therefore, to approximate the solution accurately, we need to *resolve* the spatial scale ε, using a fine computational mesh of spacing $h_f < \varepsilon$. This is not always computationally feasible, since it would require a mesh with many orders of magnitude more elements than can be handled on the world's largest supercomputers. Instead we consider four multiscale numerical techniques, as follows.

1. Homogenization and upscaling [Bensoussan, Lions, and Papanicolaou, 1978; Sanchez-Palencia, 1980]. We replace the coefficient a_ε in the differential equation by one that is easier to resolve.
2. Multiscale finite elements [Babuška and Osborn, 1983; Babuška, Caloz, and Osborn 1994; Hou and Wu 1997; Chen and Hou 2003]. We define the finite element space to better capture the fine scales.
3. Variational multiscale method [Hughes, 1995; Arbogast, Minkoff, and Keenan, 1998; Arbogast, 2000; Arbogast and Boyd, 2006]. We modify the variational form to better capture the fine scales.
4. Domain decomposition and mortar methods [Schwarz, 1870; Arbogast, Pencheva, Wheeler, and Yotov 2007]. We divide the problem into weakly coupled small subdomains that can be resolved.

2 Homogenization and Upscaling

We want to solve the problem on a coarse grid. *Upscaling* is the process of representing the system on a coarser scale by defining average or *effective macroscopic* parameters in place of the true parameters (in our case, a_ε). Perhaps the most well-developed mathematical theory of upscaling is *homogenization* [19, 62, 51, 45]. We begin with an overview of the philosophy of homogenization.

The solution $\mathbf{u}$ has high frequency "wiggles" due to the heterogeneity of a_ε, as illustrated in Fig. 1 Can we find the smooth "local average" $\bar{\mathbf{u}}(x)$ without knowing $\mathbf{u}(x)$? The wiggles are irregular, so they are hard to deal with.

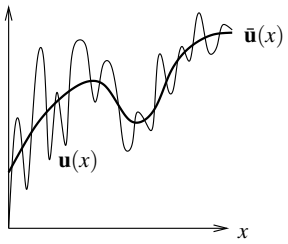


Fig. 1 The solution $\mathbf{u}$ with high frequency wiggles compared to its "local average" $\bar{\mathbf{u}}$. Can we find $\bar{\mathbf{u}}$ without knowing $\mathbf{u}$?

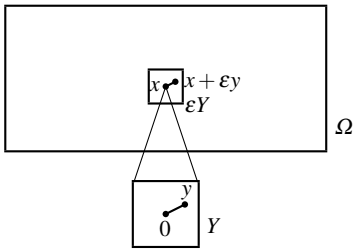


Fig. 2 The domain Ω composed of ε-scaled reference parallelepipeds Y. General location $x \in \Omega$ is modified by $y \in Y$ to give position $x+\varepsilon y$. Scaling by $1/\varepsilon$ focuses in on the details.

The key assumption in homogenization theory is that the heterogeneity has *scale separation,* meaning the system separates into fine and coarse scales with some gap in scales; that is, Fig. 1 is somehow an accurate picture of the scales separating into fine wiggles and some coarse average. More precisely, we assume the heterogeneity is *periodic* of period ε, so that the wiggles are regular, and thus easily identified and removed. Later we will see that this is basically the *closure assumption* that allows the fine-scale details to be removed uniquely from the problem. We will let $\varepsilon \to 0$, which should remove the wiggles (at least in some weak sense) and give us our macro-scale model for the average flow.

Let Y be a unit-sized reference parallelepiped cell domain, which we scale by ε. Then Ω is composed of translated copies of εY (see Fig. 2). The permeability a_ε is assumed to be, more generally, *locally periodic,* meaning that

$$a_\varepsilon(x) = a(x,y), \tag{8}$$

where $a(x,y)$ is periodic in $y \in Y$ but varies slowly and smoothly in $x \in \Omega$.

Homogenization is very mathematical, and involves deep analysis of partial differential equations [51]. Fortunately there is a simpler, more physical view of homogenization, called *formal homogenization* [19, 62].

As depicted in Fig. 2, we represent the idea of scale separation by assuming that the space variable $x_{\text{absolute}} \in \Omega$ has both a slow (denoted by $x \in \Omega$) and a fast (denoted by $y \in Y$) component, where these scale as

$$x_{\text{absolute}} \sim x + \varepsilon y.$$

At any point x_{absolute}, x is the macroscopic position that ignores fine scales and y allows us to "see" the local details under dilation by $1/\varepsilon$. In this way, we can

quantify how these local details affect larger scales. The local details disappear as $\varepsilon \to 0$, but not necessarily their coarse-scale effects.

Formal Assumption. *Assume without proof that the true solution $p(x)$ can be expanded into a power series involving ε as*

$$p(x) \sim p_0(x,y) + \varepsilon\, p_1(x,y) + \varepsilon^2 p_2(x,y) + \cdots,$$

wherein $y = x/\varepsilon$ and each p_k is periodic in $y \in Y$.

Note that the gradient operator scales as

$$\nabla \sim \nabla_x + \varepsilon^{-1}\nabla_y.$$

We expect that

$$p = p_\varepsilon \to p_0 \quad \text{as } \varepsilon \to 0.$$

Substitute the formal expansion into the equations (1)–(2) to obtain

$$-(\varepsilon^{-1}\nabla_y + \nabla_x)\cdot a(x,y)(\varepsilon^{-1}\nabla_y + \nabla_x)\sum_{k=0}^{\infty} \varepsilon^k\, p_k(x,y) = f.$$

Equating terms with like powers of ε leads to the following conclusions (for more details, see, e.g., [19, 62, 45]).

1. The ε^{-2} terms and periodicity in y imply that $p_0(x,y) = p_0(x)$ only (i.e., homogenization removes y, as we had hoped).
2. The ε^{-1} terms imply the existence and form of a *closure operator,* which is

$$p_1(x,y) := \sum_{j=1}^{d} \omega_j(x,y)\, \frac{\partial p_0(x)}{\partial x_j},$$

where the ω_j solve the local *cell problems,* one for each coordinate direction $j = 1, \ldots, d$,

$$-\nabla_y \cdot \left[a(x,y)\nabla_y \omega_j(x,y)\right] = \nabla_y \cdot \left[a(x,y)\mathbf{e}_j\right] \quad \text{in } \Omega \times Y, \tag{9}$$

$$\omega_j(x,y) \text{ is periodic in } y, \tag{10}$$

where $\mathbf{e}_j$ is the standard unit vector in the jth direction.

3. By averaging over the cell Y, the ε^0 terms give the *homogenized equations*

$$\mathbf{u}_0 = -a_0 \nabla p_0 \quad \text{in } \Omega \quad \text{(Homogenized Darcy's law)}, \tag{11}$$

$$\nabla \cdot \mathbf{u}_0 = f \quad \text{in } \Omega \quad \text{(Conservation)}, \tag{12}$$

$$\mathbf{u}_0 \cdot \nu = 0 \quad \text{on } \partial\Omega, \tag{13}$$

wherein $a_0(x)$ can be computed as the tensor

$$a_{0,ij}(x) := \frac{1}{|Y|}\int_Y a(x,y)\Big(\frac{\partial \omega_j}{\partial y_i}(x,y)+\delta_{ij}\Big)\,dy. \tag{14}$$

Lemma 1. *The homogenized permeability a_0 is symmetric and positive definite.*

Therefore a_0 has d principle eigenvectors and only positive eigenvalues, which is required of a permeability tensor in Darcy's Law (11) on physical grounds.

Lemma 2. [Voigt-Reiss Inequality] *The homogenized permeability a_0 lies between the harmonic and arithmetic averages. More precisely, if*

$$\hat{a} := \Big(\frac{1}{|Y|}\int_Y (a(x,y))^{-1}\,dy\Big)^{-1} \quad \textit{and} \quad \bar{a} := \frac{1}{|Y|}\int_Y a(x,y)\,dy,$$

then

$$\xi^T\hat{a}(x)\xi \le \xi^T a_0(x)\xi \le \xi^T\bar{a}(x)\xi \quad \forall\,\xi\in\mathbb{R}^d,\ a.e.\ x\in\Omega.$$

Thus we have the homogenized permeability tensor $a_0(x)$ from (14), and we can compute $p_0(x)$ from (11)–(13), which is well-posed by Theorem 1 and the remarks following. In fact, one can prove the following theorem on convergence [51, 8, 7], essentially justifying the first two terms in the formal asymptotic expansion.

Theorem 2. *Let p_ε and $\mathbf{u}_\varepsilon$ solve* (4)–(5)*, with a_ε satisfying the local periodicity condition* (8)*, and let p_0 and $\mathbf{u}_0$ solve* (11)–(13)*. If $p_0\in H^2(\Omega)\cap W^{1,\infty}(\Omega)$ and the* first order corrector *is defined as*

$$p_\varepsilon^1 := p_0 + \varepsilon\sum_{j=1}^d \omega_j(x,x/\varepsilon)\,\frac{\partial p_0(x)}{\partial x_j} = p_0(x)+\varepsilon\,p_1(x,x/\varepsilon),$$

then there is $C>0$, independent of ε, such that

$$\|p_\varepsilon - p_0\|_0 \le C\varepsilon \quad \textit{and} \quad \|\nabla(p_\varepsilon - p_\varepsilon^1)\|_0 \le C\sqrt{\varepsilon}.$$

Moreover, let $\alpha_0 = a_0^{-1}$ and define the fixed tensor $\mathscr{A} := a\,(I+D\omega)\alpha_0$, i.e.,

$$\mathscr{A}_{ij}(x,y) := \sum_{k,\ell} a_{ik}(x,y)\Big(\delta_{k\ell}+\frac{\partial\omega_\ell(x,y)}{\partial y_k}\Big)\alpha_{0,\ell j},$$

which is independent of ε and the domain Ω. If $\mathscr{A}_\varepsilon(x)=\mathscr{A}(x,x/\varepsilon)$, then

$$\mathbf{u}_\varepsilon(x) = \mathscr{A}_\varepsilon(x)\,\mathbf{u}_0(x) + \theta_\varepsilon^\Omega(x),$$

where

$$\|\theta_\varepsilon^\Omega\|_0 \le C\big\{\varepsilon\|\mathbf{u}_0\|_1 + \sqrt{\varepsilon|\partial\Omega|}\,\|\mathbf{u}_0\|_{0,\infty}\big\} = \mathscr{O}\big(\sqrt{\varepsilon}\,\big).$$

Herein, $\|\cdot\|_{k,p,\omega}$ denotes the norm in the Sobolev space $W^{k,p}(\omega)$. We will omit p when it is 2 and ω when it is Ω. The theory of homogenization has seemingly solved our problem with heterogeneity, since we are dealing with the case where ε is small

and so the error in Theorem 2 is negligible. However, there are serious limitations to this approach, especially in the mixed context where we are more concerned with accurate approximation of $\mathbf{u} = \mathbf{u}_\varepsilon$.

First, p_0 is approximated coarsely, and so has no microstructure. Thus

$$\mathbf{u}_0 = -a_0 \nabla p_0 \not\approx \mathbf{u}_\varepsilon.$$

We therefore need to use $p_\varepsilon^1 \approx p_\varepsilon$, which does contain the microstructure. However, even though

$$\mathbf{u}_\varepsilon^1 = -a_\varepsilon \nabla p_\varepsilon^1 \approx \mathbf{u}_\varepsilon \quad \text{and} \quad \mathscr{A}_\varepsilon(x)\,\mathbf{u}_0(x) \approx \mathbf{u}_\varepsilon,$$

we lose the divergence property, since

$$\nabla \cdot \mathbf{u}_\varepsilon^1 \neq \nabla \cdot \mathbf{u}_\varepsilon = f \quad \text{and} \quad \nabla \cdot \mathscr{A}_\varepsilon \mathbf{u}_0 \neq \nabla \cdot \mathbf{u}_\varepsilon = f.$$

This means that the local conservation principle is not satisfied. In fact, the error is $\mathscr{O}(1)$, so we are not even approximately mass conservative.

Second, a subtle question in the two-scale separation case arises: given $a_\varepsilon(x)$, what is $a(x,y)$? In practice, one works on a coarse computational grid, and, given $x \in \Omega$, one treats Y as a portion of the mesh (one or more coarse elements) around x, and sets $a(x,y) = a_\varepsilon(y)$ there. But it is not completely clear that this is appropriate.

Finally, and most importantly, we really want to develop techniques that apply to the non-two-scale separation cases. We thus turn to multiscale numerical techniques. However, we will use homogenization theory as a guide for the general case, since the two-scale separation case is the only one we completely understand.

3 Multiscale Numerics

Within the multiscale numerical approach, the objective is to solve the problem in a way that:

1. does not fully incorporate the problem dynamics (i.e., solves some global coarse scale problem to resolution $h > \varepsilon$);
2. yet captures significant features of the solution, by taking into account the microstructure (to resolution $h_f < \varepsilon$).

Many techniques fall into this general class of methods. We note the following techniques and give some references, including what we believe are the first works in the area, although our list is very incomplete, since there is a vast amount of work in the area of multiscale numerics.

1. **Multiscale finite elements** began with the work of Babuška and Osborn in 1983 and experienced major advancements in the work of Hou and Wu starting in 1997 [14, 46, 47, 35, 64]. These methods were extended to mixed systems explicitly by Chen and Hou in 2003 [28, 3, 4], but they were actually defined implicitly

for mixed systems much earlier as variational multiscale methods by Arbogast, Minkoff, and Keenan in 1998, as noted by Arbogast and Boyd in 2006 [10, 8].
2. **Variational multiscale analysis** began with the work of Hughes in 1995 [48, 49, 23], and was defined for the mixed case by Arbogast, Minkoff, and Keenan in 1998 [10, 5, 6, 8, 52, 56, 57].
3. **Multiscale multilevel and mortar methods,** in the context of homogenization or multiscale problems, can be considered to be implicit in the work of Moulton, Dendy, and Hyman in 1998, and were further developed by Xu and Zikatanov in 2004 [55, 65, 54, 41, 53]. These were extended to the mixed case in the sense of multiscale mortar methods by Arbogast, Pencheva, Wheeler, and Yotov in 2007 [11]. A multiscale basis optimization technique was defined by Rath in 2006 [59, 60].
4. **Multiscale finite volumes and discontinuous Galerkin methods** were also developed. Multiscale finite volumes were first described by Jenny, Lee, and Tchelepi in 2003 [50, 39, 43], and multiscale discontinuous Galerkin methods were defined by Aarnes and Heimsund [1].
5. **Heterogeneous multiscale methods** were defined by E and Engquist in 2003 [32].

We discuss three of these techniques in detail herein: multiscale finite elements, the variational multiscale method, and multiscale mortar methods. Each of these take an overall multiscale strategy with four main components, as follows.

1. **Localization.** The full partial differential problem is decomposed into many small, local, coarse element subproblems (of scale $h > \varepsilon$).
2. **Fine-scale effects.** The local subproblems are given appropriate boundary conditions and solved on the fine scale $h_f < \varepsilon$ (to resolve variations in a_ε) to define a coarse scale multiscale finite element or finite volume basis.
3. **Global coarse-grid problem.** This h-scale coarse basis is used to approximate the solution globally.
4. **Fine-grid reconstruction.** The finite element basis encapsulates an h_f-scale fine representation of the solution.

Note that in these methods, the problem is fully resolved on the fine scale, but the problem is *not* fully coupled. The global problem is a reduced degree-of-freedom system. Computational efficiency comes from divide-and-conquer: small, localized subproblems are easily solved; and the coupled global problem has only a few degrees of freedom per coarse element, and so is relatively easily solved.

3.1 Nonmixed Multiscale Finite Elements

For simplicity, we introduce multiscale finite elements in the nonmixed case. Recall that the objective is to define finite elements tailored to the problem at hand to better capture the fine scales.

To make everything concrete, we begin with an example in one dimension. Consider the problem

$$-(ap')' = 0, \quad 0 < x < 1, \tag{15}$$

$$p(0) = 0 \text{ and } p(1) = 1, \tag{16}$$

where $a > 0$ is highly oscillatory, leading to an oscillatory true solution, as indicated in Fig. 3. Let $X = H_0^1(0,1) = \{w \in H^1 : w(0) = w(1) = 0\}$. Then our problem has the variational form

Find $p \in X + x$ such that

$$(ap', w') = 0 \quad \forall\, w \in X.$$

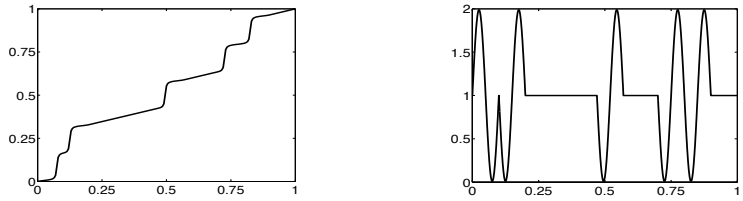


Fig. 3 The true solution p is shown on the left for the coefficient a on the right (although a becomes very small, it remains uniformly positive).

We choose a uniform grid of five points $x_i = i/4$, $i = 0,1,2,3,4$. We illustrate the definition of finite elements beginning with the standard piecewise linear basis before defining the multiscale variant.

Standard finite element space $\bar{X}_h$. At x_i, $i = 1,2,3,4$, we define $\bar{q}_i$ to be piecewise linear over the mesh such that $\bar{q}_i(x_j) = \delta_{ij}$. To state this constructively, let $x_5 = 1$. As illustrated in Fig. 4, we define $\bar{q}_i$, supported in (x_{i-1}, x_{i+1}), by:

1. Setting $\bar{q}_i$ on the boundary of each element separately, so $\bar{q}_i(x_j) = \delta_{ij}$;
2. Linearly interpolating over each element separately;
3. Joining the two pieces together continuously to form $\bar{q}_i$ and setting $\bar{X}_h = \text{span}\{\bar{q}_i\}$.

Multiscale finite element space X_h. Localize X to the element $E = (x_{i-1}, x_i)$ as $X(E) = H_0^1(E)$. As illustrated in Fig. 5, at x_i, $i = 1,2,3,4$, we define q_i, supported in (x_{i-1}, x_{i+1}), by:

1. Setting q_i on the boundary of each element separately, so $q_i(x_j) = \delta_{ij}$;
2. Solving the homogeneous problem on each element E

   Find $q_i \in X(E) + \bar{q}_i(x)$ such that

$$(aq_i', w')_E = 0 \quad \forall\, w \in X(E),$$

   where E is (x_{i-1}, x_i) or (x_i, x_{i+1}), using the appropriate linear function $\bar{q}_i(x)$ for the boundary conditions;

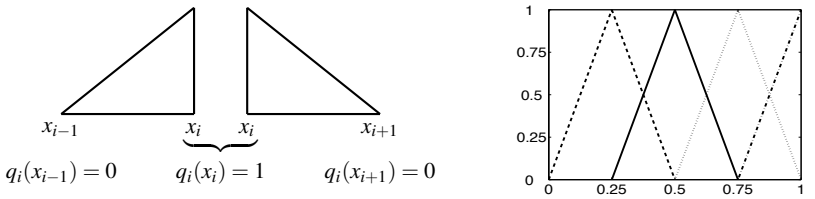


Fig. 4 Construction of standard piecewise linear basis functions in one dimension.

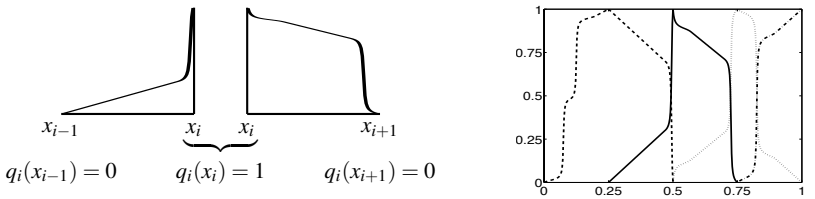


Fig. 5 Construction of multiscale basis functions in one dimension.

3. Joining the two pieces together continuously to form q_i and setting $X_h = \text{span}\{q_i\}$.

We illustrate the finite element solutions that result from using standard and multiscale finite elements in Fig. 6. This is merely an illustration, since the multiscale finite elements reproduce the exact solution in one dimension, but *not* in higher dimensions. To avoid misleading the reader, an error has been displayed. The point is that standard elements simply cannot represent the microstructure, whereas multiscale elements have this ability.

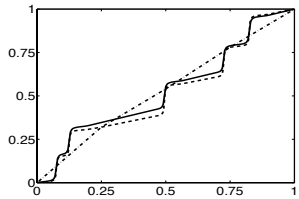


Fig. 6 The solid line is the true solution, the dashed line illustrates the multiscale finite element solution (actually, it is exact), and the dashed-dotted line is the standard finite element solution.

For the general multi-dimensional problem

$$-\nabla \cdot a_\varepsilon \nabla p = f \quad \text{in } \Omega, \tag{17}$$

$$-a_\varepsilon \nabla p \cdot \nu = 0 \quad \text{on } \partial\Omega, \tag{18}$$

the standard variational problem is

Find $p \in X = H^1(\Omega)/\mathbb{R}$ such that

$$A_\varepsilon(p,w) := (a_\varepsilon \nabla p, \nabla w) = (f,w) \quad \forall\, w \in X. \tag{19}$$

General finite element construction. Let $\mathscr{T}_h$ be a finite element partition of Ω. Define standard $\bar{q}_i$ and multiscale q_i finite elements on an element $E \in \mathscr{T}_h$ as follows.

1. Set $\bar{q}_i$ and q_i on ∂E to be some simple polynomial. More generally, in the multiscale case, we can set $q_i = \ell_i(x)$ on ∂E, where ℓ_i is any appropriate function.
2. Use some polynomial interpolation over E to define $\bar{q}_i$. However, for multiscale elements, we solve the homogeneous problem on each element E

   Find $q_i \in X(E) + \ell_i(x)$ such that

$$A_\varepsilon(q_i, w)_E = 0 \quad \forall\, w \in X(E),$$

   i.e., we solve the Dirichlet problems (on a fine grid)

$$-\nabla \cdot a_\varepsilon \nabla q_i = 0 \quad \text{in } E, \tag{20}$$

$$q_i = \ell_i \quad \text{on } \partial E. \tag{21}$$

3. Join the pieces together continuously to form $\bar{X}_h = \text{span}\{\bar{q}_i\}$ and $X_h = \text{span}\{q_i\}$.

We remark that the multiscale approach has a lot of flexibility in Steps 1 and 2, and there exist many variants of the above procedure.

Multiscale structure of X_h. When we set q_i to be $\bar{q}_i$ on each element boundary (i.e., we take $\ell_i = \bar{q}_i$), we can exhibit the multiscale structure of X_h by noting that

$$q_i = \bar{q}_i + (q_i - \bar{q}_i) =: \bar{q}_i + q_i'.$$

In this form, we define q_i' by

Find $q_i' \in X(E)$ such that

$$A_\varepsilon(q_i', w')_E = -A_\varepsilon(\bar{q}_i, w')_E \quad \forall\, w' \in X(E).$$

The $q_i' \in X(E) = H_0^1(E)$ are "bubble functions," localized to a coarse element by imposing homogeneous Dirichlet boundary conditions. They are fine-scale and contain the microstructure information. The $\bar{q}_i$ are coarse-scale.

Theorem 3. *Let $X_h' = span\{q_i'\}$. Then*

$$X_h = span\{\bar{q}_i + q_i'\} \subsetneq \bar{X}_h \oplus X_h'$$

is a Hilbert space direct sum decomposition into coarse and fine scales.

3.2 Nonmixed Variational Multiscale Method

Again for simplicity, we introduce the variational multiscale method in the nonmixed case, treating the system (17)–(18). Recall that the objective is to modify the variational form of the differential system (19) to better capture the fine scales.

We begin by separating the solution space $X = H^1(\Omega)/\mathbb{R}$ into coarse and fine scales using a Hilbert space direct sum decomposition. Let

$$X = \bar{X} \oplus X', \tag{22}$$

and separate the standard variational form into coarse and fine scales through the test functions as

Find $p = \bar{p} + p' \in \bar{X} \oplus X'$ such that

$$A_\varepsilon(\bar{p} + p', \bar{w}) = (f, \bar{w}) \quad \forall\, \bar{w} \in \bar{X} \quad \text{(Coarse scales)}, \tag{23}$$
$$A_\varepsilon(\bar{p} + p', w') = (f, w') \quad \forall\, w' \in X' \quad \text{(Fine scales)}. \tag{24}$$

Rewrite the fine scale equation as

$$A_\varepsilon(p', w') = (f, w') - A_\varepsilon(\bar{p}, w') \quad \forall\, w' \in X',$$

and note that this is a well defined problem for p'. It implicitly defines an *affine* upscaling operator taking $\bar{X}$ to X'. The linear part of the operator is $\hat{p}' : \bar{X} \to X'$, and it satisfies

$$A_\varepsilon(\hat{p}'(\bar{q}), w') = -A_\varepsilon(\bar{q}, w') \quad \forall\, w' \in X'. \tag{25}$$

The constant part of the upscaling operator is $\tilde{p}' \in X'$, which satisfies

$$A_\varepsilon(\tilde{q}', w') = (f, w') \quad \forall\, w' \in X'. \tag{26}$$

The full upscaling operator is $\hat{p}'(\cdot) + \tilde{p}' : \bar{X} \to X'$, and given coarse scales, we can obtain fine scales as

$$p' = \hat{p}'(\bar{p}) + \tilde{p}'. \tag{27}$$

Now the coarse scale equation is simply

$$A_\varepsilon\big(\bar{p} + \hat{p}'(\bar{p}), \bar{w}\big) = (f, \bar{w}) - A_\varepsilon(\tilde{p}', \bar{w}) \quad \forall\, \bar{w} \in \bar{X},$$

and the effect of the fine scales is manifest within this coarse-scale variational problem. Taking $w' = \hat{p}'(\bar{w})$ in (25) enables us to symmetrize the form to

$$A_\varepsilon\big(\bar{p} + \hat{p}'(\bar{p}), \bar{w} + \hat{p}'(\bar{w})\big) = (f, \bar{w}) - A_\varepsilon(\tilde{p}', \bar{w}) \quad \forall\, \bar{w} \in \bar{X}. \tag{28}$$

If we define the bilinear and linear forms to be

$$B_\varepsilon(\bar{p}, \bar{w}\big) = A_\varepsilon\big(\bar{p} + \hat{p}'(\bar{p}), \bar{w} + \hat{p}'(\bar{w})\big) \quad \text{and} \quad \mathscr{F}(\bar{w}) = (f, \bar{w}) - A_\varepsilon(\tilde{p}', \bar{w}),$$

then we have the modified variational form

$$B_\varepsilon(\bar{p}, \bar{w}\big) = \mathscr{F}(\bar{w}) \quad \forall\, \bar{w} \in \bar{X}. \tag{29}$$

In the variational multiscale method, both the bilinear and linear forms are modified.

Choice of Hilbert space decomposition. To be useful for finite element approximation, we need to localize the fine scales. For $\mathscr{T}_h$ a (coarse) finite element partition of Ω, let

$$X' := \bigoplus_E X(E) = \bigoplus_E H_0^1(E),$$

and then $X = \bar{X} \oplus X'$, where

$$\bar{X} = X/X' \simeq \{q|_e : e \text{ is a coarse edge of } \mathscr{T}_h\}.$$

Thus $\bar{X}$ is determined by its values on $\partial E \; \forall \, E \in \mathscr{T}_h$.

Finite element approximation. We use a standard Galerkin finite element space $\bar{X}_h = \{\bar{q}_h\}$ and the multiscale fine space X'. That is, X' is localized and

$$\bar{X}_h \oplus X' \subsetneq \bar{X} \oplus X' = H^1/\mathbb{R}.$$

In practice, we must further approximate $X'|_E = X(E)$ on each element E by a finite element space on a very fine mesh $X_{h_f}(E)$. However, since E is small, we can make this approximation as accurate as we need, and so for simplicity we will assume that it is handled exactly.

We have three equivalent ways to describe the finite element approximation. The primary method is given by direct approximation of (29); however, it is instructive to instead start from the original two-scale decomposition (23)–(24). This leads to

Version 1. Find $p_h = \bar{p}_h + p'_h \in \bar{X}_h \oplus X'$ such that

$$A_\varepsilon(p_h, w) = (f, w) \quad \forall \, w \in \bar{X}_h \oplus X'.$$

But $\bar{X}_h \oplus X'$ is a very large space. In fact, $\bar{p}_h$ and p'_h are related, and the solution is in a much smaller space.

Since Galerkin methods minimize energy, the multiscale solution minimizes energy in the large space $\bar{X}_h \oplus X'$. For these methods, if one specifies the value of the finite elements on ∂E, then the best approximation comes from using the finite element that minimizes energy within E.

Theorem 4. *If the multiscale finite elements are specified on ∂E for each element $E \in \mathscr{T}_h$, then the best approximation comes from using the multiscale finite element that minimizes energy within E.*

By solving for the upscaling operator as above, we obtain the equivalent form

Version 2. Find $\bar{p}_h \in \bar{X}_h$ such that

$$B_\varepsilon(\bar{p}_h, \bar{w}) = \mathscr{F}(\bar{w}) \quad \forall \, \bar{w} \in \bar{X}_h.$$

Now $\bar{X}_h$ is very small, but we must find the upscaling operator to relate $\bar{q}_h$ and $p'_h(\bar{q}_h)$. Given a basis $\bar{X}_h = \text{span}\{\bar{q}_i\}$, we solve a local Dirichlet problem for each $\bar{q}_i$ on element E

$$A_\varepsilon(\bar{q}_i + \hat{p}'(\bar{q}_i), w')_E = 0 \quad \forall \, w' \in X(E).$$

These are the same problems as in the multiscale finite element case, so $X_h = \operatorname{span}\{\bar{q}_i + \hat{p}'(\bar{q}_i)\}$ are the same elements from Theorem 3, and we can reformulate the variational multiscale method as a multiscale finite element method

Version 3. Find $p_h \in X_h$ such that

$$A_\varepsilon(p_h, w) = (f, w) - A_\varepsilon(\tilde{p}', w) \quad \forall\, w \in X_h.$$

Theorem 5. *Up to treatment of f (i.e., $\tilde{p}'$), the variational multiscale and multiscale finite element methods are the same in this basic setting.*

Unlike multiscale finite elements, the variational multiscale method naturally handles nonzero f. Henceforth we will use this correction in the multiscale finite element method as well.

4 Mixed Variational Multiscale Method

The mixed case (4)–(5) is complicated by the fact that we treat directly both the scalar unknown $p \in W = L^2(\Omega)/\mathbb{R}$ and the vector unknown $\mathbf{u} \in \mathbf{V} = H_0(\operatorname{div};\Omega)$. We base our two-scale expansion of the solution space $W \times \mathbf{V}$ on the local mass conservation principle. Given a coarse computational mesh of elements $\mathscr{T}_h$ on Ω with element edges (or faces) $\mathscr{E}_h$, let the *pressure space* be $W = \bar{W} \oplus W'$, where

$$\begin{aligned} \bar{W} &\subset \{\bar{w} \in W : \bar{w} \text{ is constant on each coarse element } E \in \mathscr{T}_h\}, \\ W' &:= \bar{W}^\perp. \end{aligned}$$

The *velocity space* is then $\mathbf{V} = \bar{\mathbf{V}} \oplus \mathbf{V}'$, where

$$\begin{aligned} \mathbf{V}' &:= \{\mathbf{v}' \in \mathbf{V} : \nabla \cdot \mathbf{v}' \in W',\ \mathbf{v}' \cdot \nu = 0 \text{ on } \partial E\ \forall\, E \in \mathscr{T}_h\}, \\ \bar{\mathbf{V}} &:= \mathbf{V}/\mathbf{V}' \simeq \{\mathbf{v} \cdot \nu \text{ on } \partial E : E \in \mathscr{T}_h\}. \end{aligned}$$

Note that $\mathbf{V}'$ is localized by imposing homogeneous Neumann boundary conditions, leaving $\bar{\mathbf{V}}$ with full normal velocity coupling on the coarse edges $e \in \mathscr{E}_h$. Note also that we have decomposed $\mathbf{V}$ according to coarse and fine scales related to mass conservation, since we can define $\bar{\mathbf{V}}$ explicitly in such a way that

$$\begin{aligned} \nabla \cdot \bar{\mathbf{V}} &= \bar{W} \quad \text{(coarse conservation)}, \\ \nabla \cdot \mathbf{V}' &= W' \quad \text{(fine subgrid conservation)}. \end{aligned}$$

We can now separate scales uniquely via the direct sum as

$$p = \bar{p} + p' \in \bar{W} \oplus W' \quad \text{and} \quad \mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}' \in \bar{\mathbf{V}} \oplus \mathbf{V}'.$$

Moreover, we can separate the variational form (4)–(5) into coarse scales

$$(a_\varepsilon^{-1}(\bar{\mathbf{u}}+\mathbf{u}'),\bar{\mathbf{v}})-(\bar{p},\nabla\cdot\bar{\mathbf{v}})=0 \quad \forall\,\bar{\mathbf{v}}\in\bar{\mathbf{V}}, \tag{30}$$

$$(\nabla\cdot\bar{\mathbf{u}},\bar{w})=(f,\bar{w}) \quad \forall\,\bar{w}\in\bar{W}, \tag{31}$$

and fine scales

$$(a_\varepsilon^{-1}(\bar{\mathbf{u}}+\mathbf{u}'),\mathbf{v}')-(p',\nabla\cdot\mathbf{v}')=0 \quad \forall\,\mathbf{v}'\in\mathbf{V}', \tag{32}$$

$$(\nabla\cdot\mathbf{u}',w')=(f,w') \quad \forall\,w'\in W'. \tag{33}$$

As noted in [6], we have the following well-posedness result.

Lemma 3. *The inf-sup condition holds over both $\bar{W}\times\bar{\mathbf{V}}$ and $W'\times\mathbf{V}'$, with constants independent of the coarse mesh and ε. Moreover, given $\bar{\mathbf{u}}\in\bar{\mathbf{V}}$, there exists a unique solution $(p',\mathbf{u}')\in W'\times\mathbf{V}'$ and*

$$\|p'\|_0+\|\mathbf{u}'\|_{\mathbf{V}}\le C\{\|f\|_0+\|\bar{\mathbf{u}}\|_0\}.$$

Thus we can define a *closure operator* relating fine scales to coarse scales from (32)–(33). This is an affine operator, with linear part defined for $\bar{\mathbf{v}}\in\bar{\mathbf{V}}$ as $(\hat{p}',\hat{\mathbf{u}}')\in W'\times\mathbf{V}'$, where

$$(a_\varepsilon^{-1}(\bar{\mathbf{v}}+\hat{\mathbf{u}}'),\mathbf{v}')-(\hat{p}',\nabla\cdot\mathbf{v}')=0 \quad \forall\,\mathbf{v}'\in\mathbf{V}', \tag{34}$$

$$(\nabla\cdot\hat{\mathbf{u}}',w')=0 \quad \forall\,w'\in W', \tag{35}$$

and constant part defined as $(\tilde{p}',\tilde{\mathbf{u}}')\in W'\times\mathbf{V}'$, where

$$(a_\varepsilon^{-1}\tilde{\mathbf{u}}',\mathbf{v}')-(\tilde{p}',\nabla\cdot\mathbf{v}')=0 \quad \forall\,\mathbf{v}'\in\mathbf{V}', \tag{36}$$

$$(\nabla\cdot\tilde{\mathbf{u}}',w')=(f,w') \quad \forall\,w'\in W'. \tag{37}$$

That is,

$$p'=\hat{p}'(\bar{\mathbf{u}})+\tilde{p}' \quad \text{and} \quad \mathbf{u}'=\hat{\mathbf{u}}'(\bar{\mathbf{u}})+\tilde{\mathbf{u}}'.$$

Lemma 4. *The operator $\hat{\mathbf{u}}':\bar{\mathbf{V}}\to\mathbf{V}'\cap\ker(\nabla\cdot)$ is bounded and linear.*

Using the upscaling operator to replace fine-scale quantities in (32)–(33), we obtain the upscaled variational problem (written in symmetric form)

Find $(\bar{p},\bar{\mathbf{u}})\in\bar{W}\times\bar{\mathbf{V}}$ such that

$$(a_\varepsilon^{-1}(\bar{\mathbf{u}}+\hat{\mathbf{u}}'(\bar{\mathbf{u}})),(\bar{\mathbf{v}}+\hat{\mathbf{u}}'(\bar{\mathbf{v}})))-(\bar{p},\nabla\cdot\bar{\mathbf{v}})=-(a_\varepsilon^{-1}\tilde{\mathbf{u}}',\bar{\mathbf{v}}) \quad \forall\,\bar{\mathbf{v}}\in\bar{\mathbf{V}}, \tag{38}$$

$$(\nabla\cdot\bar{\mathbf{u}},\bar{w})=(f,\bar{w}) \quad \forall\,\bar{w}\in\bar{W}, \tag{39}$$

with the full solution given by

$$p=\bar{p}+\hat{p}'(\bar{\mathbf{u}})+\tilde{p}' \quad \text{and} \quad \mathbf{u}=\bar{\mathbf{u}}+\hat{\mathbf{u}}'(\bar{\mathbf{u}})+\tilde{\mathbf{u}}'.$$

Note that the equations maintain strict local conservation on both scales.

We can also rewrite the problem as

Find $(\bar{p},\bar{\mathbf{u}})\in\bar{W}\times\bar{\mathbf{V}}$ such that

$$(a_\varepsilon^{-1}\bar{\mathbf{u}},\bar{\mathbf{v}}) - (a_\varepsilon^{-1}\hat{\mathbf{u}}'(\bar{\mathbf{u}}),\hat{\mathbf{u}}'(\bar{\mathbf{v}})) - (\bar{p},\nabla\cdot\bar{\mathbf{v}}) = -(a_\varepsilon^{-1}\tilde{\mathbf{u}}',\bar{\mathbf{v}}) \quad \forall\,\bar{\mathbf{v}}\in\bar{\mathbf{V}},$$
$$(\nabla\cdot\bar{\mathbf{u}},\bar{w}) = (f,\bar{w}) \quad \forall\,\bar{w}\in\bar{W}.$$

The positive diffusion term $(a_\varepsilon^{-1}\bar{\mathbf{u}},\bar{\mathbf{v}})$ has a negative subscale correction, which is therefore *antidiffusive* on the coarse scale. This is the main reason that effective parameters (merely replacing a_ε on E by some average quantity in the original equations) cannot work well. Rather, some multiscale ideas are needed.

Numerical Approximation. Choose any inf-sup stable mixed space $\bar{W}_h\times\bar{\mathbf{V}}_h$ on the coarse mesh [26, 21]. Then we approximate $p\approx p_h$ and $\mathbf{u}\approx\mathbf{u}_h$ as

Version 1. Find $(\bar{p}_h,\bar{\mathbf{u}}_h)\in\bar{W}_h\times\bar{\mathbf{V}}_h$ such that

$$\big(a_\varepsilon^{-1}(\bar{\mathbf{u}}_h+\hat{\mathbf{u}}'(\bar{\mathbf{u}}_h)),\bar{\mathbf{v}}_h+\hat{\mathbf{u}}'(\bar{\mathbf{v}}_h)\big) - (\bar{p}_h,\nabla\cdot\bar{\mathbf{v}}_h) = -(a_\varepsilon^{-1}\tilde{\mathbf{u}}',\bar{\mathbf{v}}_h) \quad \forall\,\bar{\mathbf{v}}_h\in\bar{\mathbf{V}}_h, \quad (40)$$
$$(\nabla\cdot\bar{\mathbf{u}}_h,\bar{w}_h) = (f,\bar{w}_h) \quad \forall\,\bar{w}_h\in\bar{W}_h, \quad (41)$$

and then set

$$p_h = \bar{p}_h + \hat{p}'(\bar{\mathbf{u}}_h) + \tilde{p}' \quad\text{and}\quad \mathbf{u}_h = \bar{\mathbf{u}}_h + \hat{\mathbf{u}}'(\bar{\mathbf{u}}_h) + \tilde{\mathbf{u}}'. \quad (42)$$

By defining the space

$$\mathbf{V}_h := \{\bar{\mathbf{v}}_h+\hat{\mathbf{u}}'(\bar{\mathbf{v}}_h)\,:\,\bar{\mathbf{v}}_h\in\bar{\mathbf{V}}_h\}\subsetneq\bar{\mathbf{V}}_h+\mathbf{V}', \quad (43)$$

we can express the method as the multiscale finite element method

Version 2. Find $\bar{p}_h\in\bar{W}_h$ and $\mathbf{u}_h\in\mathbf{V}_h+\tilde{\mathbf{u}}'$ such that

$$(a_\varepsilon^{-1}\mathbf{u}_h,\mathbf{v}_h) - (\bar{p}_h,\nabla\cdot\mathbf{v}_h) = 0 \quad \forall\,\mathbf{v}_h\in\mathbf{V}_h, \quad (44)$$
$$(\nabla\cdot\mathbf{u}_h,\bar{w}_h) = (f,\bar{w}_h) \quad \forall\,\bar{w}_h\in\bar{W}_h, \quad (45)$$

with the reconstruction $p_h=\bar{p}_h+\hat{p}'(\bar{\mathbf{u}}_h)+\tilde{p}'$. Furthermore, after some manipulation of the equations, we can express the method in the non-computable form

Version 3. Find $p_h\in W$ and $\mathbf{u}_h\in\bar{V}_h+\mathbf{V}'$ such that

$$(a_\varepsilon^{-1}\mathbf{u}_h,\mathbf{v}_h) - (p_h,\nabla\cdot\mathbf{v}_h) = 0 \quad \forall\,\mathbf{v}_h\in\bar{\mathbf{V}}_h+\mathbf{V}', \quad (46)$$
$$(\nabla\cdot\mathbf{u}_h,w) = (f,w) \quad \forall\,w\in W. \quad (47)$$

5 Mixed Multiscale Finite Elements

The mixed multiscale finite element method is defined above in (44)–(45). Our task is to define the discrete spaces. We will define some of the simplest mixed multiscale finite elements that are commonly used. In all cases, we will take the *pressure space*

$$\bar{W}_h := \{\bar{w}\in L^2(\Omega)\,:\,\bar{w}\text{ is constant on each coarse element }E\in\mathscr{T}_h\}.$$

We deal with the fact that $W_h\not\subset W=L^2(\Omega)/\mathbb{R}$ in the usual way. Since

$$\bar{\mathbf{V}} \simeq \{\mathbf{v}\cdot\nu \text{ on } \partial E : E \in \mathscr{T}_h\},$$

we need only specify $\bar{\mathbf{v}} \in \bar{\mathbf{V}}$ on coarse element edges $e \in \mathscr{E}_h$. We obtain the corresponding multiscale finite element $\mathbf{v}_h$ by solving the local Neumann problem. For simplicity, we work in two dimensions on rectangular elements.

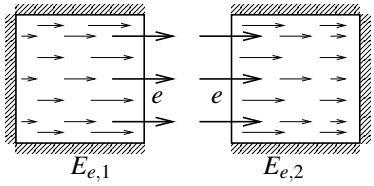
Raviart-Thomas mixed elements (RT0). The standard lowest order Raviart-Thomas vector variable space $\mathbf{V}_h^{\mathrm{RT0}}$ [61] has one basis function $\mathbf{v}_e^{\mathrm{RT0}}$ associated with each coarse element edge $e \in \mathscr{E}_h$. The degrees of freedom are the constants $\mathbf{v}\cdot\nu|_e$ for each edge $e \in \mathscr{E}_h$. For example, if $E = [0,1]^2$ then we have the four pieces

$$\mathbf{v}_{\mathrm{L}} = \begin{pmatrix} 1-x \\ 0 \end{pmatrix}, \quad \mathbf{v}_{\mathrm{R}} = \begin{pmatrix} x \\ 0 \end{pmatrix}, \quad \mathbf{v}_{\mathrm{B}} = \begin{pmatrix} 0 \\ 1-y \end{pmatrix}, \quad \mathbf{v}_{\mathrm{T}} = \begin{pmatrix} 0 \\ y \end{pmatrix},$$

which are joined to neighbors across the edge for which $\mathbf{v}\cdot\nu = 1$. This is the standard polynomial definition.

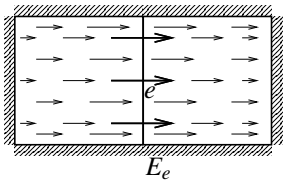
We can also define these finite elements as the solution to two types of differential equations. The first we will call the *element definition.* For each edge $e \in \mathscr{E}_h$, let $E_{e,i}$, $i = 1,2$, be the two elements that contain e. We solve on $E = E_{e,i}$, $i = 1,2$,

$$\begin{cases} \mathbf{v}_e^{\mathrm{RT0}} = -\nabla\phi_e^{\mathrm{RT0}} & \text{in } E, \\ \nabla\cdot\mathbf{v}_e^{\mathrm{RT0}} = \pm|e|/|E| & \text{in } E, \\ \mathbf{v}_e^{\mathrm{RT0}}\cdot\nu = \begin{cases} 0 & \text{on } \partial E\setminus e, \\ 1 & \text{on } e. \end{cases} \end{cases} \tag{48}$$



However, it is equally valid to define $\mathbf{v}_e^{\mathrm{RT0}}$ on the *dual-support* element $E_e = E_{e,1} \cup E_{e,2}$ by solving

$$\begin{cases} \mathbf{v}_e^{\mathrm{RT0}} = -\nabla\phi_e^{\mathrm{RT0}} & \text{in } E_e, \\ \nabla\cdot\mathbf{v}_e^{\mathrm{RT0}} = \pm|e|/|E_{e,i}| & \text{in } E_{e,i},\ i = 1,2, \\ \mathbf{v}_e^{\mathrm{RT0}}\cdot\nu = 0 & \text{on } \partial E_e. \end{cases} \tag{49}$$



We have the following convergence result [61]. Since these elements have no dependence on the scale ε, they are accurate only when $h < \varepsilon$, i.e., h resolves the fine-scale heterogeneity.

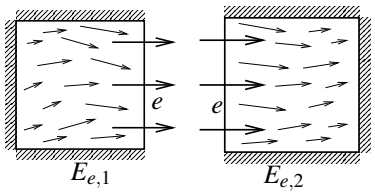
Theorem 6. $\|\mathbf{u} - \mathbf{u}_h^{RT0}\|_0 \le C\|\mathbf{u}\|_1 h = \mathscr{O}(h/\varepsilon)$.

The main idea of multiscale finite elements is to use a_ε in their definition. In the boundary value problems above, we simply insert the coefficient a_ε.

Variational multiscale element (ME0) based on RT0. The simplest multiscale element (ME0) is due implicitly to Arbogast, Minkoff, and Keenan in 1998 [10] (cf. [8, 28]). It is based on using RT0 as the coarse space in the variational multiscale

method, or, equivalently, using the element definition of the RT0 element above. Define $\mathbf{v}_e^{\mathrm{ME0}} \in \mathbf{V}_h^{\mathrm{ME0}}$ for each coarse element edge $e \in \mathscr{E}_h$ by solving

$$\begin{cases} \mathbf{v}_e^{\mathrm{ME0}} = -a_\varepsilon \nabla \phi_e^{\mathrm{ME0}} & \text{in } E, \\ \nabla \cdot \mathbf{v}_e^{\mathrm{ME0}} = \pm |e|/|E| & \text{in } E, \\ \mathbf{v}_e^{\mathrm{ME0}} \cdot \nu = \begin{cases} 0 & \text{on } \partial E \setminus e, \\ 1 & \text{on } e. \end{cases} \end{cases} \tag{50}$$



We have the following convergence result [6, 28, 8].

Theorem 7. *In general,* $\|\mathbf{u} - \mathbf{u}_h^{ME0}\|_0 \le C\|\mathbf{u}\|_1 h = \mathscr{O}(h/\varepsilon)$. *In the two-scale separation case,* $\mathbf{u}_0$ *is the solution of the homogenized problem and*

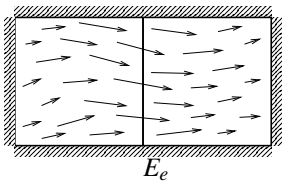
$$\|\mathbf{u} - \mathbf{u}_h^{ME0}\|_0 \le C\Big\{ \|\mathbf{u}_0\|_1 h + \|\mathbf{u}_0\|_0 \varepsilon + \|\mathbf{u}_0\|_{0,\infty} \sqrt{\varepsilon/h} \Big\}.$$

Since $\mathbf{u}_0$ *is independent of* ε,

$$\|\mathbf{u} - \mathbf{u}_h^{ME0}\|_0 = \mathscr{O}\big(\min\big\{ h/\varepsilon \ , \ h + \varepsilon + \sqrt{\varepsilon/h} \big\} \big).$$

Multiscale dual-support (MD) elements. Elements based on RT0 can also be defined using the dual-support definition. This was done first by Aarnes et al. [3, 4]. Define $\mathbf{v}_e^{\mathrm{MD}} \in \mathbf{V}_h^{\mathrm{MD}}$ for each coarse element edge $e \in \mathscr{E}_h$ by solving

$$\begin{cases} \mathbf{v}_e^{\mathrm{MD}} = -a_\varepsilon \nabla \phi_e^{\mathrm{MD}} & \text{in } E_e, \\ \nabla \cdot \mathbf{v}_e^{\mathrm{MD}} = \pm |e|/|E_{e,i}| & \text{in } E_{e,i},\ i = 1,2, \\ \mathbf{v}_e^{\mathrm{MD}} \cdot \nu = 0 & \text{on } \partial E_e. \end{cases} \tag{51}$$



The shape on $E_{e,1}$ depends on $E_{e,2}$, and vice-versa. Thus, as defined by Ciarlet [31], this is not a finite element. Nevertheless, we will consider it to be a finite element. It has a problem with convergence in the classical sense (i.e., the error should vanish as $h \to 0$) [7]. As noted in [7], the problem is related to anisotropy. For example, if one takes a constant

$$a_\varepsilon(x) = a = Q\Lambda Q^T \quad \text{with } \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ and } Q \text{ a rotation,}$$

we have a genuine anisotropy when the rotation is not a multiple of $\pi/2$ and $\lambda_1 \ne \lambda_2$, but no microstructure. On $e \in \mathscr{E}_h$, the MD element $\mathbf{v}_e^{\mathrm{MD}}$ has a nonconstant normal trace $\mathbf{v}_e^{\mathrm{MD}} \cdot \nu$. Therefore the space $\mathbf{V}_h^{\mathrm{MD}}$ cannot reproduce constants, so the method cannot converge in any reasonable sense as $h \to 0$.

It should be noted that these elements are designed to be effective when $h > \varepsilon$. If for some reason one would take $h \to 0$, one should also change elements when h

becomes smaller than $\mathscr{O}(\varepsilon)$. Moreover, there are many variants of this basic element as defined herein which improve the convergence properties when $h > \varepsilon$.

Second order accurate elements (BDM1 and ME1). The standard elements of Brezzi, Douglas, and Marini, defined in 1985, also use a piecewise constant scalar space $\bar{W}_h$, but they are formally second order accurate for the velocity space $\mathbf{V}_h^{\mathrm{BDM1}}$ [25, 24]. The BDM1 elements have two degrees of freedom per element edge. That is,

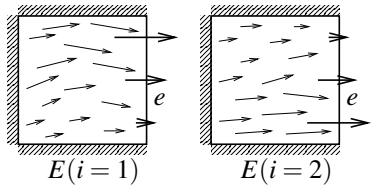
$$\mathbf{v}^{\mathrm{BDM1}} \cdot \nu|_e \text{ is a linear function for each edge } e \in \mathscr{E}_h.$$

However, we maintain the conservation property that $\nabla \cdot \mathbf{v}|_E$ is a constant on each element $E \in \mathscr{T}_h$ for all $\mathbf{v}^{\mathrm{BDM1}} \in \mathbf{V}_h^{\mathrm{BDM1}}$, so $\nabla \cdot \mathbf{V}_h^{\mathrm{BDM1}} = \bar{W}_h$. More precisely, on a rectangular element $E \in \mathscr{T}_h$, a finite element in $\mathbf{V}_h^{\mathrm{BDM1}}$ has eight degrees of freedom $a, b, \ldots, h$ as

$$\mathbf{v}^{\mathrm{BDM1}} = \begin{pmatrix} a + bx + cy + 2gxy - hx^2 \\ d + ex + fy - gy^2 + 2hxy \end{pmatrix}.$$

The multiscale variant, due to Arbogast in 2000 [5], is defined using BDM1 as the coarse space in the variational multiscale method, giving $\mathbf{V}_h^{\mathrm{ME1}}$ with two degrees of freedom per edge $e \in \mathscr{E}_h$. For $i = 1, 2$ and L_i a basis for linear polynomials on e, we construct $\mathbf{v}_{e,i}^{\mathrm{ME1}}$ by solving

$$\begin{cases} \mathbf{v}_{e,i}^{\mathrm{ME1}} = -a_\varepsilon \nabla \phi_{e,i}^{\mathrm{ME1}} & \text{in } E, \\ \nabla \cdot \mathbf{v}_{e,i}^{\mathrm{ME1}} = \dfrac{1}{|E|} \displaystyle\int_e L_i & \text{in } E, \\ \mathbf{v}_{e,i}^{\mathrm{ME1}} \cdot \nu = \begin{cases} 0 & \text{on } \partial E \setminus e, \\ L_i & \text{on } e, \end{cases} \end{cases} \qquad (52)$$



and joining the two pieces from each side of e.

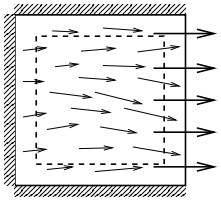
We have the following convergence result [6, 8].

Theorem 8. *In general,* $\|\mathbf{u} - \mathbf{u}_h^{ME1}\|_0 \le C\|\mathbf{u}\|_2 h^2 = \mathscr{O}(h^2/\varepsilon^2)$. *In the two-scale separation case,* $\mathbf{u}_0$ *is the solution of the homogenized problem and*

$$\|\mathbf{u} - \mathbf{u}_h^{ME1}\|_0 \le C\Big\{ \|\mathbf{u}_0\|_2 h^2 + \|\mathbf{u}_0\|_0 \varepsilon + \|\mathbf{u}_0\|_{0,\infty} \sqrt{\varepsilon/h} \Big\} = \mathscr{O}\big(h^2 + \varepsilon + \sqrt{\varepsilon/h}\big).$$

Some additional elements. As noted earlier, the multiscale finite element approach allows great flexibility in their definition. We note here four main variants.

1. **Oversampled elements (OS).** Hou et al. [46, 28] proposed an *oversampling* technique for the local partial differential systems used to define multiscale finite elements. Instead of solving on $E \in \mathscr{T}_h$, one solves on a larger domain, and then restricts the solution back to E. This gives a nonconforming method, since the pieces do not join continuously across edges. The advantage of oversampling is that it increases the variability on ∂E and allows for better multiscale approximation.



2. **Generalized finite elements and partition of unity methods.** Babuška et al. [16, 14, 64] advocate creating a multiscale finite element basis from local multiscale functions, perhaps defined by oversampling as above. However, instead of simply restricting back to the element $E \in \mathscr{T}_h$ (or E_e for $e \in \mathscr{E}_h$), one uses a partition of unity method so that the resulting elements are conforming.
3. **Reduced dimension-based elements.** Hou and Wu [46] proposed defining the multiscale boundary condition on $e \in \mathscr{E}_h$ by solving a reduced dimension problem. In the case of the nonmixed system, one would first solve the lower dimensional problem $-\nabla_e \cdot a_\varepsilon \nabla_e p_e = 0$ on e for p_e. Then one sets the boundary condition in (20)–(21) to be $\ell_i = p_e$. Again this improves the variability on ∂E. It is not so clear that this technique applies to the mixed problem, since we need information normal to e (i.e., $\mathbf{v} \cdot \nu$).
4. **Local eigenfunction-based elements.** Efendiev, Galvis, and Wu, as well as Hetmaniuk and Lehoucq [33, 44] propose that, in the differential problems defining the multiscale finite elements, the boundary condition $\mathbf{v} \cdot \nu$ on $e \in \mathscr{E}_h$ be based on the solution of a local eigenfunction problem (solved in, e.g., E_e and restricted to e). The eigenbasis is the most efficient basis, and so should give a good definition of the local boundary condition on e. The energy minimizing extension into $E_{e,i}$, $i = 1,2$, is then the best choice for finite element, as noted above in Theorem 4. Techniques for efficient definition of the multiscale basis, and for reducing its dimension, are given in [33].
5. **Homogenization-based elements (HE).** Arbogast [7] proposes using homogenization theory to define $\mathbf{v} \cdot \nu$ on $e \in \mathscr{E}_h$ and energy minimizing extension into $E_{e,i}$, $i = 1,2$. The idea is easily seen from the homogenization theorem, Theorem 2. The microstructure is

$$\mathbf{u}_\varepsilon \approx \mathscr{A}_\varepsilon \mathbf{u}_0,$$

that is, $\mathbf{u}_\varepsilon$ is a fixed, ε-scale multiple of a smooth function. Since we know how to approximate a smooth function (e.g., by a polynomial), we should define

$$\mathbf{V}_h \approx \{\mathscr{A}_\varepsilon \mathbf{v} : \mathbf{v} \text{ is some nice smooth function}\}.$$

However, these finite elements lie outside $H(\mathrm{div};\Omega)$, so we use the idea merely to define the normal velocity. For $e \in \mathscr{E}_h$, we approximate the smooth part, corresponding to $\mathbf{u}_0 \cdot \nu|_e$, by a constant vector. Thus we have two degrees of freedom per edge, and a basis for the *homogenization-based finite element space* (HE) $\mathbf{V}_h^{\mathrm{HE}}$ can be constructed by energy minimizing extension of the normal traces

$$\mathscr{A}_\varepsilon \mathbf{e}_i \cdot \nu|_e \quad \text{to} \quad \mathbf{v}_{e,i}^{\mathrm{HE}}, \quad i = 1,2.$$

6 A Multiscale Analysis of the Approximation Errors

In this section, we analyze the variational multiscale method (40)–(41), or, equivalently, the multiscale finite element method (44)–(45) with the $\tilde{\mathbf{u}}'$ correction term. Throughout this section, we will tacitly assume that a_ε is uniformly positive definite (7) and that $\nabla \cdot \bar{\mathbf{V}}_h \subset \bar{W}_h$. We noted already that the MD finite elements do not converge as $h \to 0$, but the ME0 and ME1 elements are better behaved. We assume that the upscaling operator (i.e., the local problems on E) is solved exactly, since it can be well resolved on a fine grid. Because our problems are well-posed (recall Lemma 3), small discretization errors on the subgrid scale will propagate boundedly in the error estimates that we present later. However, for completeness, we note in passing the following theorem, which accounts for subgrid approximation for standard multiscale elements [6].

Theorem 9. *In the variational multiscale method* (40)–(41)*, suppose that inf-sup stable finite elements $\bar{W}_H \times \bar{\mathbf{V}}_H$ are used on the coarse scale H which approximate velocity to order L. Suppose also that the upscaling operator is approximated on a subgrid of element size $h < H$ by a mixed method $W_h \times \mathbf{V}_h$ approximating the velocity to order ℓ. If $\mathscr{P}_{W_h}$ denotes fine grid L^2-projection onto $W_h = \bar{W}_h + W_h'$, then*

$$\|a_\varepsilon^{-1/2}(\mathbf{u} - \mathbf{u}_h)\|_0 \le \inf_{\substack{\mathbf{v}_h \in \bar{\mathbf{V}}_H + \mathbf{V}_h' \\ \nabla \cdot \mathbf{v}_h = \mathscr{P}_{W_h} f}} \|a_\varepsilon^{-1/2}(\mathbf{u} - \mathbf{v}_h)\|_0 \le C_\varepsilon H^L,$$

$$\nabla \cdot \mathbf{u}_h = \mathscr{P}_{W_h} f.$$

Moreover, if pressure is approximated to coarse order M and fine order $m \le M$, then

$$\|\mathscr{P}_{W_h} p - p_h\|_0 \le C_\varepsilon (H^{M-m} h^{m+1} + H^{L+1}),$$
$$\|p - p_h\|_0 \le C_\varepsilon (H^{M-m} h^m + H^{L+1}).$$

Of course, the constants suffer from the problem of scale. However, in the resolved case, the combination of coarse BDM1 spaces (i.e., ME1) with fine RT0 gives the nice estimates ($L = 2$, $l = M = m = 1$)

$$\|\mathbf{u} - \mathbf{u}_h\|_0 \le C_\varepsilon H^2, \quad \|\nabla \cdot (\mathbf{u} - \mathbf{u}_h)\|_0 \le Ch, \quad \text{and} \quad \|p - p_h\|_0 \le C_\varepsilon (h + H^3),$$

which suggests the scaling $H = \sqrt{h}$, giving $\mathscr{O}(h)$ convergence.

We turn now to multiscale error analysis that quantifies the error in terms of h and ε. The proof is based on comparison to the homogenized solution, and so applies technically only to the two-scale separation case, i.e., when (8) holds, which we tacitly assume throughout this section. The style of proof is due to Hou et al. [47, 35]

in groundbreaking work on the multiscale analysis of finite element methods. For the mixed case, see also [28, 8].

We begin with a general quasioptimality result [6].

Theorem 10. *For any* $\mathbf{v} \in \bar{\mathbf{V}}_h + \mathbf{V}'$ *such that* $\nabla \cdot \mathbf{v} = \nabla \cdot \mathbf{u}_\varepsilon = f$,

$$\|a_\varepsilon^{-1/2}(\mathbf{u}_\varepsilon - \mathbf{u}_h)\|_0 \le \|a_\varepsilon^{-1/2}(\mathbf{u}_\varepsilon - \mathbf{v})\|_0,$$
$$\nabla \cdot \mathbf{u}_h = f.$$

Proof. Since we assumed the upscaling operator is solved exactly, we obtain precisely that $\nabla \cdot \mathbf{u}_h = f$ from (47). This means that the multiscale method is locally conservative on the fully resolved fine scale.

The velocity error is bounded by subtracting (4) and (46) and taking a test function $\mathbf{v} - \mathbf{u}_h$, where $\mathbf{v} \in \bar{\mathbf{V}}_h + \mathbf{V}'$ has the required divergence. That we optimize over the larger space $\bar{\mathbf{V}}_h + \mathbf{V}'$ instead of $\mathbf{V}_h$ is a consequence of the fact that the upscaling operator is defined as energy minimizing extension. □

We sketch now a simplified multiscale convergence proof [7] involving certain projection operators and four key results. We analyze here only ME0 and avoid complexities like oversampling, thus proving the multiscale part of Theorem 7. The first key result is quasioptimality, Theorem 10. The second key result is homogenization, Theorem 2, which says that $\mathbf{u}_\varepsilon \approx \mathscr{A}_\varepsilon \mathbf{u}_0$. Thus our goal is to find any $\mathbf{v}_\varepsilon \approx \mathscr{A}_\varepsilon \mathbf{u}_0$ in $\mathbf{V}_h^{\mathrm{ME0}} + \tilde{\mathbf{u}}'$ with $\nabla \cdot \mathbf{v}_\varepsilon = \nabla \cdot \mathbf{u}_\varepsilon = f$.

The third key result involves dealing with the ε scale of our finite elements, so we define corresponding *homogenized finite elements.* We replace the true coefficient a_ε in the definition of the finite elements (50) with the corresponding homogenized one a_0, giving a finite element space

$$\mathbf{V}_{0,h}^{\mathrm{ME0}} = \operatorname*{span}_{e \in \mathscr{E}_h} \{\mathbf{v}_{0,e}^{\mathrm{ME0}}\}.$$

Since our finite elements are defined by boundary value problems, the homogenization theorem applies, although we will see numerical resonance (i.e., factors of ε/h) in the estimate, which come from localizing to the element E_e of scale h.

Lemma 5. *For each* $e \in \mathscr{E}_h$,

$$\mathbf{v}_e^{ME0} = \mathscr{A}_\varepsilon \mathbf{v}_{0,e}^{ME0} + \theta_\varepsilon^{E_e,ME0},$$
$$\|\theta_\varepsilon^{E_e,ME0}\|_{0,E_e} \le C\big\{\varepsilon \|\mathbf{v}_{0,e}^{ME0}\|_{1,E_e} + \sqrt{\varepsilon |\partial E_e|}\, \|\mathbf{v}_{0,e}^{ME0}\|_{0,\infty,E_e}\big\} = \mathscr{O}\Big(\frac{\varepsilon}{h} + \sqrt{\frac{\varepsilon}{h}}\Big) h^{d/2}.$$

We next define flux-based projection operators for both $\mathbf{V}_h^{\mathrm{ME0}}$ and $\mathbf{V}_{0,h}^{\mathrm{ME0}}$, which are related through the lemma above. The average normal flux across $e \in \mathscr{E}_h$ is

$$\gamma_e = \frac{1}{|e|} \int_e \mathbf{v} \cdot \nu_e \, ds.$$

The usual Raviart-Thomas projection is defined as

$$\pi^{\mathrm{RT0}}\mathbf{v} := \sum_{e\in\mathscr{E}_h} \gamma_e \mathbf{v}_e^{\mathrm{RT0}} \in \mathbf{V}_h^{\mathrm{RT0}}.$$

Similarly, we define

$$\pi_\varepsilon^{\mathrm{ME0}}\mathbf{v} := \sum_{e\in\mathscr{E}_h} \gamma_e \mathbf{v}_e^{\mathrm{ME0}} \in \mathbf{V}_h^{\mathrm{ME0}} \quad \text{and} \quad \pi_0^{\mathrm{ME0}}\mathbf{v} := \sum_{e\in\mathscr{E}_h} \gamma_e \mathbf{v}_{0,e}^{\mathrm{ME0}} \in \mathbf{V}_{0,h}^{\mathrm{ME0}},$$

leading us to the third key result.

Lemma 6. *Let $\mathscr{P}_{\bar{W}_h}$ be L^2-projection onto $\bar{W}_h$, the space of piecewise discontinuous constants. Then*

$$\nabla\cdot\pi_\varepsilon^{ME0}\mathbf{v} = \nabla\cdot\pi_0^{ME0}\mathbf{v} = \nabla\cdot\pi^{RT0}\mathbf{v} = \mathscr{P}_{\bar{W}_h}\nabla\cdot\mathbf{v},$$
$$\|\pi_\varepsilon^{ME0}\mathbf{v} - \mathscr{A}_\varepsilon\pi_0^{ME0}\mathbf{v}\|_0 \le C\|\mathbf{v}\|_1\Big(\varepsilon/h + \sqrt{\varepsilon/h}\Big).$$

Proof. The divergence condition follows from the Divergence Theorem. For the estimate, note that

$$\pi_\varepsilon^{\mathrm{ME0}}\mathbf{v} - \mathscr{A}_\varepsilon\pi_0^{\mathrm{ME0}}\mathbf{v} = \sum_{e\in\mathscr{E}_h}\gamma_e(\mathbf{v}_e^{\mathrm{ME0}} - \mathscr{A}_\varepsilon\mathbf{v}_{0,e}^{\mathrm{ME0}}) = \sum_{e\in\mathscr{E}_h}\gamma_e\theta_e^{E_e,\mathrm{ME0}}.$$

Theorem 2 on homogenization gives us that

$$\begin{aligned}
\|\pi_\varepsilon^{\mathrm{ME0}}\mathbf{v} - \mathscr{A}_\varepsilon\pi_0^{\mathrm{ME0}}\mathbf{v}\|_{0,E} &\le \sum_{e\subset\partial E}|\gamma_e|\,\|\theta_e^{E_e,\mathrm{ME0}}\|_{0,E} \\
&\le C\sum_{e\subset\partial E}\Big(h^{-d/2}\|\mathbf{v}\|_{1,E_e}\Big)\Big(\frac{\varepsilon}{h}+\sqrt{\frac{\varepsilon}{h}}\Big)h^{d/2} \\
&= C\sum_{e\subset\partial E}\|\mathbf{v}\|_{1,E_e}\Big(\frac{\varepsilon}{h}+\sqrt{\frac{\varepsilon}{h}}\Big),
\end{aligned}$$

and the proof is completed by squaring, summing over all $E\in\mathscr{T}_h$, and noting that there is finite overlap of the E_e. □

The fourth and final key result concerns smooth approximation by π_0^{ME0}. In general, this is not a nice operator. However, for the types of vector fields we consider, it approximates well.

Lemma 7. *If $\mathbf{v}_0 = -a_0\nabla\phi_0$, then*

$$\|\mathbf{v}_0 - \pi_0^{ME0}\mathbf{v}_0\|_0 \le C\|\mathbf{v}_0\|_1 h.$$

Proof. Let

$$\psi = \mathbf{v} - \pi_0^{\mathrm{ME0}}\mathbf{v} = -a_0\nabla\Big(\phi_0 - \sum_{e\subset\partial E}\gamma_e\phi_{0,e}^{\mathrm{ME0}}\Big) \quad \text{in } E,$$

which is a potential field satisfying the Neumann problem

$$\begin{aligned}\nabla\cdot\psi &= \nabla\cdot\mathbf{v}_0 - \mathscr{P}_{\bar{W}_h}\nabla\cdot\mathbf{v}_0 \quad \text{in } E,\\ \psi\cdot\nu_e &= \mathbf{v}_0\cdot\nu_e - \gamma_e \qquad\qquad \text{on } e\subset\partial E.\end{aligned}$$

The standard energy estimate (see, e.g., [36, 38]) gives the result. □

We are now ready to state and prove the discrete inf-sup condition. We need to use the elliptic regularity theorem (see, e.g., [36, 38, 21]), which requires that Ω have, e.g., a $C^{1,1}$ boundary or that Ω be convex [42].

Lemma 8. *If Ω supports elliptic regularity, then there is some $\beta>0$, independent of ε, such that*

$$\sup_{\mathbf{v}_h\in\mathbf{V}_h^{ME0}} \frac{(\bar{w}_h,\nabla\cdot\mathbf{v}_h)}{\|\mathbf{v}_h\|_0+\|\nabla\cdot\mathbf{v}_h\|_0} \geq \beta\|\bar{w}_h\|_0 \quad \forall\ \bar{w}_h\in\bar{W}_h.$$

Proof. Recall that $\bar{w}_h$ is orthogonal to constants. Solve

$$\begin{aligned}\nabla\cdot\mathbf{v}_0 &= \bar{w}_h \qquad &&\text{in } \Omega, &(53)\\ \mathbf{v}_0 &= -a_0\nabla\phi_0 \quad &&\text{in } \Omega, &(54)\\ \mathbf{v}_0\cdot\nu &= 0 \qquad &&\text{on } \partial\Omega, &(55)\end{aligned}$$

and note that elliptic regularity implies that

$$\|\mathbf{v}_0\|_1 \leq C\|\bar{w}_h\|_0.$$

Take

$$\mathbf{v}_h = \pi_\varepsilon^{\mathrm{ME0}}\mathbf{v}_0 \in \mathbf{V}_h^{\mathrm{ME0}},$$

for which

$$\nabla\cdot\mathbf{v}_h = \mathscr{P}_{\bar{W}_h}\nabla\cdot\mathbf{v}_0 = \bar{w}_h.$$

Then, by the third and fourth key results,

$$\begin{aligned}\|\mathbf{v}_h\|_0 &\leq \|\pi_\varepsilon^{\mathrm{ME0}}\mathbf{v}_0 - \mathscr{A}_\varepsilon\pi_0^{\mathrm{ME0}}\mathbf{v}_0\|_0 + \|\mathscr{A}_\varepsilon(\pi_0^{\mathrm{ME0}}\mathbf{v}_0-\mathbf{v}_0)\|_0 + \|\mathscr{A}_\varepsilon\mathbf{v}_0\|_0\\ &\leq C\|\mathbf{v}_0\|_1 \leq C\|\bar{w}_h\|_0,\end{aligned}$$

and the result follows. □

Theorem 11. *If Ω supports elliptic regularity and $p_0\in H^2(\Omega)\cap W^{1,\infty}(\Omega)$, then*

$$\begin{aligned}&\|\mathbf{u}_\varepsilon-\mathbf{u}_h^{ME0}\|_0 + \|\mathscr{P}_{\bar{W}_h}p_\varepsilon-\bar{p}_h\|_0\\ &\qquad\leq C\big\{\big(\varepsilon+\varepsilon/h+\sqrt{\varepsilon/h}+h\big)\|\mathbf{u}_0\|_1+\sqrt{\varepsilon}\,\|\mathbf{u}_0\|_{0,\infty}\big\},\\ &\nabla\cdot\mathbf{u}_h^{ME0} = f.\end{aligned}$$

Proof. Since $\mathbf{u}_\varepsilon\approx\pi_\varepsilon^{\mathrm{ME0}}\mathbf{u}_0+\tilde{\mathbf{u}}'\in\mathbf{V}_h^{\mathrm{ME0}}+\tilde{\mathbf{u}}'$ and $\nabla\cdot(\pi_\varepsilon^{\mathrm{ME0}}\mathbf{u}_0+\tilde{\mathbf{u}}') = \mathscr{P}_{\bar{W}_h}\nabla\cdot\mathbf{u}_0+\nabla\cdot\tilde{\mathbf{u}}' = f$, by quasioptimality, key result one or Theorem 10, we have that

$$\begin{aligned}\|\mathbf{u}_\varepsilon - \mathbf{u}_h^{\mathrm{ME0}}\|_0 &\le C\|\mathbf{u}_\varepsilon - \pi_\varepsilon^{\mathrm{ME0}}\mathbf{u}_0 - \tilde{\mathbf{u}}'\|_0 \\ &\le C\big\{\|\mathbf{u}_\varepsilon - \mathscr{A}_\varepsilon\mathbf{u}_0\|_0 + \|\mathscr{A}_\varepsilon(\mathbf{u}_0 - \pi_0^{\mathrm{ME0}}\mathbf{u}_0)\|_0 \\ &\qquad + \|\mathscr{A}_\varepsilon\pi_0^{\mathrm{ME0}}\mathbf{u}_0 - \pi_\varepsilon^{\mathrm{ME0}}\mathbf{u}_0\|_0 + \|\tilde{\mathbf{u}}'\|_0\big\},\end{aligned}$$

and the velocity estimate follows from the final three key results, Theorem 2, Lemmas 6 and 7, and the standard energy estimate of (36)–(37), which says that

$$\|\tilde{\mathbf{u}}'\|_0 \le C\|\mathscr{P}_{\tilde{W}_h}^{\perp} f\|_{-1} \le C\|f\|_0 h \le C\|\mathbf{u}_0\|_1 h.$$

The divergence result follows in general, and the pressure result follows from the inf-sup condition, Lemma 8 and the difference of (4) and (46). □

We remark that a similar proof holds for ME1 [7]. We also note that recent work by Babuška and Lipton [15] and Efendiev, Galvis, and Wu [33] shows multiscale convergence for certain multiscale methods independently of the two-scale separation hypothesis (8).

7 Domain Decomposition and Mortar Methods

In this section, we discuss a restricted class of domain decomposition and mortar methods related to the mixed finite element methods considered earlier for our heterogeneous elliptic problem. In 1988, Glowinski and Wheeler [40] defined nonoverlapping domain decomposition for mixed methods by iterating on the Dirichlet-to-Neumann map. As depicted in Fig. 7, given the pressure λ on the subdomain interfaces Γ, one computes the flow locally. Based on the flux mismatch on Γ (i.e., the jump in $\mathbf{u}\cdot\nu$), one updates λ using, e.g., conjugate gradients. Once converged, the full fine-scale problem is solved. The technique allows great flexibility in handling interdomain multiphysics (different physical models in different subdomains) and is well suited to parallel computation. It allows us to handle interdomain multiscale aspects as well.

In 1994, Bernardi, Maday, and Patera [20] defined mortar methods to glue the subdomains together weakly when the subdomain meshes do not match. As illustrated in Fig. 8, Arbogast, Cowsar, Wheeler, and Yotov in 2000 [9] extended the mortar idea to mixed methods, using a continuous or discontinuous linear mortar λ. The idea was to use grid spacings of $\mathscr{O}(h)$ for all grids.

The mixed mortar method is similar to our previous multiscale techniques. It has the same four basic components noted in Sec. 3 above: localization to subdomains, fine-scale effects resolved on the subdomains, a global interface problem for the mortar unknowns, and fine-grid reconstruction over Ω. If the mortar resolves the computational meshes, i.e., the subdomain and mortar mesh spacings are $\mathscr{O}(h)$, $h < \varepsilon$, we obtain a fully resolved and weakly but fully coupled approximation.

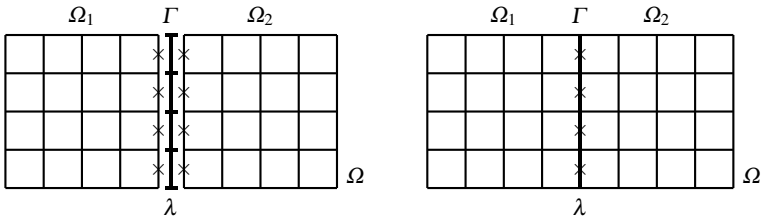


Fig. 7 Illustration of domain decomposition. The domain Ω on the right is shown separated on the left for clarity. On the interface Γ, λ resolves the computational mesh on both $\partial\Omega_1$ and $\partial\Omega_2$.

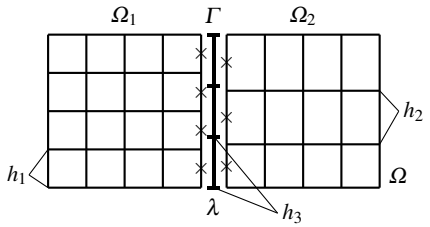


Fig. 8 Illustration of mortar mixed methods. On the interface Γ, λ does not match the computational mesh on $\partial\Omega_1$ and/or $\partial\Omega_2$. The idea here is to consider that h_1, h_2, and h_3 are of the same order, so the problem is fully coupled.

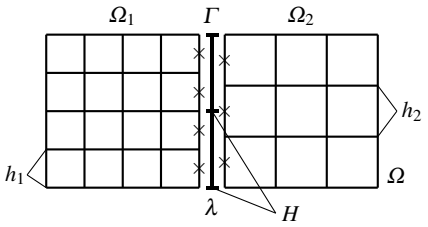


Fig. 9 Illustration of multiscale mortar mixed methods. Here, H is much coarser than h_1 and h_2. We use a higher order mortar approximation to compensate for the coarseness of the grid and maintain good overall accuracy.

The idea behind the multiscale mortar mixed method is to relax the coupling dynamics as in multiscale methods. As illustrated in Fig. 9, we use the following four main components.

1. **Localization.** Divide Ω into many small subdomains (or coarse elements of scale H), over which the original partial differential system is imposed.
2. **Fine-scale effects.** The subdomains are given Dirichlet boundary conditions $p = \lambda$ on Γ and solved on the fine scale h to define the local solution.
3. **Global coarse-grid interface problem.** The weakly defined flux mismatch (jump in $\mathbf{u}\cdot\nu$) on Γ is used to define a better λ on scale $H > h$, and we iterate this and the previous step until convergence is attained.
4. **Fine-grid representation of the solution.** We obtain a fully resolved and well coupled approximate solution if λ is approximated in a higher order space.

To be more precise, let Ω be decomposed into nonoverlapping subdomains Ω_i, which correspond to coarse elements in our previous methods. Define the interfaces

$$\Gamma_{ij} := \partial\Omega_i \cap \partial\Omega_j, \quad \Gamma := \bigcup_{i<j} \Gamma_{ij}, \quad \text{and} \quad \Gamma_i := \partial\Omega_i \cap \Gamma.$$

With ν_i denoting the outer unit normal to $\partial\Omega_i$, the differential problem (1)–(3) is equivalent to the decomposed system

$$a^{-1}\mathbf{u} = -\nabla p \quad \text{in } \Omega_i \quad \text{(subdomain Darcy's law)}, \tag{56}$$
$$\nabla \cdot \mathbf{u} = f \quad \text{in } \Omega_i \quad \text{(subdomain conservation)}, \tag{57}$$
$$\mathbf{u}|_{\Omega_i} \cdot \nu_i + \mathbf{u}|_{\Omega_j} \cdot \nu_j = 0 \quad \text{on } \Gamma_{ij} \quad \text{(conservation on interface } \Gamma), \tag{58}$$
$$p|_{\Omega_i} = p|_{\Omega_j} \quad \text{on } \Gamma_{ij} \quad \text{(continuity of } p \text{ on } \Gamma), \tag{59}$$
$$\mathbf{u} \cdot \nu = 0 \quad \text{on } \partial\Omega. \tag{60}$$

A variational form is

Find $p \in L^2(\Omega_i)$, $\mathbf{u} \in H_0(\mathrm{div};\Omega_i)$, and $\lambda = p \in H^{1/2}(\Gamma)$ such that

$$(a^{-1}\mathbf{u},\mathbf{v})_{\Omega_i} - (p,\nabla\cdot\mathbf{v})_{\Omega_i} + \langle\lambda,\mathbf{v}\cdot\nu_i\rangle_{\Gamma_i} = 0 \quad \forall \mathbf{v} \in H_0(\mathrm{div};\Omega_i), \tag{61}$$
$$(\nabla\cdot\mathbf{u},w)_{\Omega_i} = (f,w)_{\Omega_i} \quad \forall w \in L^2(\Omega_i), \tag{62}$$
$$\sum_i \langle\mathbf{u}\cdot\nu_i,\mu\rangle_{\Gamma_i} = 0 \quad \forall \mu \in H^{1/2}(\Gamma), \tag{63}$$

where we use $\langle\cdot,\cdot\rangle$ for interface inner products for emphasis and $H_0(\mathrm{div};\Omega_i) = \{\mathbf{v} \in H(\mathrm{div};\Omega_i) : \mathbf{v}\cdot\nu = 0 \text{ on } \partial\Omega\}$. The last equation enforces continuity of flux on Γ.

The multiscale mortar mixed method. For finite element approximation, define on each Ω_i a fine scale finite element partition $\mathscr{T}_h^{\Omega_i}$ of maximal element diameter $h < \varepsilon$, and let $W_{h,i} \times \mathbf{V}_{h,i} \subset L^2(\Omega_i) \times H_0(\mathrm{div};\Omega_i)$ be any of the usual mixed finite element spaces. For the mortar, define a coarse scale finite element partition $\mathscr{T}_H^{\Gamma_{ij}}$ on each Γ_{ij} of maximal element diameter H, and let $M_{H,ij}$ be a space of continuous or discontinuous finite elements. For $W_h = \cup_i W_{h,i}$, $\mathbf{V}_h = \cup_i \mathbf{V}_{h,i}$, and $M_H = \cup_{ij} M_{H,ij}$, the multiscale mortar method is

Find $p_h \in W_h$, $\mathbf{u}_h \in \mathbf{V}_h$, and $\lambda_H \in M_H$ such that

$$(a^{-1}\mathbf{u}_h,\mathbf{v})_{\Omega_i} - (p_h,\nabla\cdot\mathbf{v})_{\Omega_i} + \langle\lambda_H,\mathbf{v}\cdot\nu_i\rangle_{\Gamma_i} = 0 \quad \forall \mathbf{v} \in \mathbf{V}_{h,i}, \tag{64}$$
$$(\nabla\cdot\mathbf{u}_h,w)_{\Omega_i} = (f,w)_{\Omega_i} \quad \forall w \in W_{h,i}, \tag{65}$$
$$\sum_i \langle\mathbf{u}_h\cdot\nu_i,\mu\rangle_{\Gamma_i} = 0 \quad \forall \mu \in M_H. \tag{66}$$

Now the last equation enforces only *weak* continuity of flux on Γ.

The usual way to solve this system is to reduce it to an interface problem [40]. Since it is an affine problem in λ, we define the bilinear and linear forms on M_H by

$$d_H(\lambda,\mu) := \sum_i d_{H,i}(\lambda,\mu) = -\sum_i \langle\hat{\mathbf{u}}_h(\lambda)\cdot\nu_i,\mu\rangle_{\Gamma_i},$$
$$g_H(\mu) := \sum_i g_{H,i}(\mu) = \sum_i \langle\tilde{\mathbf{u}}_h\cdot\nu_i,\mu\rangle_{\Gamma_i},$$

where $(\hat{\mathbf{u}}_h(\lambda),\hat{p}_h(\lambda)) \in \mathbf{V}_h \times W_h$ solves the linear part of the problem (i.e., with λ given and $f = 0$)

$$(a^{-1}\hat{\mathbf{u}}_h(\lambda),\mathbf{v})_{\Omega_i} - (\hat{p}_h(\lambda),\nabla\cdot\mathbf{v})_{\Omega_i} = -\langle\lambda,\mathbf{v}\cdot\nu_i\rangle_{\Gamma_i} \quad \forall \mathbf{v} \in \mathbf{V}_{h,i}, \tag{67}$$
$$(\nabla\cdot\hat{\mathbf{u}}_h(\lambda),w)_{\Omega_i} = 0 \quad \forall w \in W_{h,i}, \tag{68}$$

and $(\tilde{\mathbf{u}}_h, \tilde{p}_h) \in \mathbf{V}_h \times W_h$ solves the constant part (i.e., with $\lambda = 0$ and f given)

$$(a^{-1}\tilde{\mathbf{u}}_h, \mathbf{v})_{\Omega_i} - (\tilde{p}_h, \nabla\cdot\mathbf{v})_{\Omega_i} = 0 \quad \forall \mathbf{v} \in \mathbf{V}_{h,i}, \tag{69}$$

$$(\nabla\cdot\tilde{\mathbf{u}}_h, w)_{\Omega_i} = (f, w)_{\Omega_i} \quad \forall w \in W_{h,i}. \tag{70}$$

Theorem 12. *The interface bilinear form $d_H(\cdot,\cdot)$ is symmetric and positive definite on M_H. In fact,*

$$d_H(\lambda,\mu) = \big(a^{-1}\hat{\mathbf{u}}_h(\lambda), \hat{\mathbf{u}}_h(\mu)\big). \tag{71}$$

Moreover,

$$d_H(\lambda_H, \mu) = g_H(\mu) \quad \forall \mu \in M_H$$

if, and only if, the solution to (64)–(66) *satisfies*

$$p_h = \hat{p}_h(\lambda_H) + \tilde{p}_h \quad \textit{and} \quad \mathbf{u}_h = \hat{\mathbf{u}}_h(\lambda_H) + \tilde{\mathbf{u}}_h.$$

Thus, our problem reduces to a symmetric and positive definite linear system, and it can be solved, for example, by conjugate gradient iteration. In that case, the computations involve once solving for $(\tilde{\mathbf{u}}_h, \tilde{p}_h)$ to get $g_H(\mu)$, and at each iteration k, solving for $\big(\hat{\mathbf{u}}_h(\lambda_H^k), \hat{p}_h(\lambda_H^k)\big)$ to get $d_H(\lambda_H^k, \mu)$.

Multiscale finite element formulation. Implicit in the method are multiscale finite elements. To see them, let $\{\mu_\ell\}$ be a basis for M_H and define

$$w_\ell := \hat{p}_h(\mu_\ell) \quad \text{and} \quad \mathbf{v}_\ell := \hat{\mathbf{u}}_h(\mu_\ell). \tag{72}$$

Then

$$\lambda_H = \sum_\ell \lambda_\ell \mu_\ell, \quad p_h = \sum_\ell \lambda_\ell w_\ell + \tilde{p}_h, \quad \text{and} \quad \mathbf{u}_h = \sum_\ell \lambda_\ell \mathbf{v}_\ell + \tilde{\mathbf{u}}_h,$$

and the method seeks $\{\lambda_\ell\}$ such that

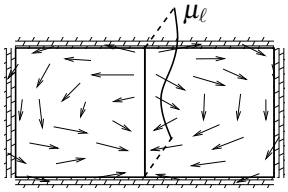
$$\sum_\ell \lambda_\ell\, d_H(\mu_\ell, \mu_k) = g_H(\mu_k) \quad \forall k,$$

which, using (71) and (70) and taking $\lambda = \mu_k$ and $\mathbf{v} = \tilde{\mathbf{u}}_h$ in (67), is equivalent to

$$\sum_\ell \lambda_\ell (a^{-1}\mathbf{v}_\ell, \mathbf{v}_k) = (f, w_k) - (a^{-1}\tilde{\mathbf{u}}_h, \mathbf{v}_k) \quad \forall k.$$

The multiscale finite elements are now evident [11, 37]. Let

$$\begin{aligned} N_{h,H} &:= \operatorname{span}\left\{ \begin{pmatrix} w_\ell \\ \mathbf{v}_\ell \end{pmatrix} \right\} \\ &= \operatorname{span}\left\{ \begin{pmatrix} \hat{p}_h(\mu_\ell) \\ \hat{\mathbf{u}}_h(\mu_\ell) \end{pmatrix} \right\} \subset \begin{pmatrix} W_h \\ \mathbf{V}_h \end{pmatrix}. \end{aligned} \tag{73}$$



Then the method is

Find $(p_h, \mathbf{u}_h) \in N_{h,H} + (\tilde{p}_h, \tilde{\mathbf{u}}_h)$ such that

$$(a^{-1}\mathbf{u}_h, \mathbf{v}) = (f, w) \quad \forall (w, \mathbf{v}) \in N_{h,H}.$$

This is an unusual multiscale finite element space. Not only do we couple pressures and velocities, but we allow nonzero normal flow on the boundary of the "coarse elements" $\partial\Omega_i$ (it is merely weakly zero). However, our multiscale finite elements are indeed locally defined over the subdomains.

Variational multiscale method formulation. We can also relate our method to the variational multiscale method in several ways, and thereby extract different sets of multiscale finite elements. If we decompose the discrete mortar space $L^2(\Gamma) = M_H \oplus M_H'$ and drop M_H', we obtain the original formulation with multiscale finite elements (73).

Another approach is to decompose the velocity space. Define the *weakly continuous normal velocities*

$$\mathbf{V}_w := \Big\{ \mathbf{v} \in \mathbf{V}_h : \sum_i \langle \mathbf{v} \cdot \nu_i, \mu \rangle_{\Gamma_i} = 0 \; \forall \mu \in M_H \Big\}.$$

The method reduces to: Find $p_h \in W_h$ and $\mathbf{u}_h \in \mathbf{V}_w$, such that

$$(a^{-1}\mathbf{u}_h, \mathbf{v}) - \sum_i (p_h, \nabla \cdot \mathbf{v})_{\Omega_i} = 0 \quad \forall \mathbf{v} \in \mathbf{V}_w, \tag{74}$$

$$\sum_i (\nabla \cdot \mathbf{u}_h, w)_{\Omega_i} = (f, w) \qquad \forall w \in W_h. \tag{75}$$

Our Hilbert space decomposition of $\mathbf{V}_w$ involves the *weakly zero normal velocities*

$$\mathbf{V}_w' := \{ \mathbf{v} \in \mathbf{V}_w : \langle \mathbf{v} \cdot \nu, \mu \rangle_{\Gamma_{ij}} = 0 \; \forall \mu \in M_{H,ij} \text{ and } \forall i, j \}.$$

Then $\bar{\mathbf{V}}_w \simeq \mathbf{V}_w / \mathbf{V}_w'$ is defined by its degrees of freedom on the interfaces as

$$\bar{\mathbf{V}}_w \simeq \big\{ \langle \mathbf{v}, \mu_\ell \rangle : \mathbf{v} \in \mathbf{V}_w \; \forall \ell \big\}. \tag{76}$$

With $W_w' := \nabla \cdot \mathbf{V}_w'$ and $\bar{W}_w := (W_w')^\perp$, we have the decomposition

$$W_h = \bar{W}_w \oplus W_w' \quad \text{and} \quad \mathbf{V}_w = \bar{\mathbf{V}}_w \oplus \mathbf{V}_w'.$$

Proceeding as before, we obtain coarse and fine scale equations analogous to (30)–(33), an upscaling operator analogous to (34)–(37), and the upscaled equation analogous to (38)–(39).

The point is that we obtain formally the same variational multiscale method as before, but now we use *nonconforming* elements with greater flexibility near $\partial\Omega_i$. The greater flexibility results from the fact that we control the normal velocities only weakly. According to (76) a basis can be found by finding for each ℓ *any* function $\mathbf{v}_{w,\ell} \in \mathbf{V}_w$ such that $\langle \mathbf{v}_{w,\ell}, \mu_k \rangle = \delta_{\ell k}$. That is, we specify $\mathbf{v}_{w,\ell} \cdot \nu$ on the boundaries of the "course elements" $\partial\Omega_i$, but only in the weak sense. Moreover, the boundary

condition on the other coarse edges $\Gamma_i \cup \Gamma_j \setminus \Gamma_{ij}$ is only weakly zero, so some flow between subdomains is allowed on the fine scale.

We have the following a-priori error estimates [11].

Theorem 13. *If p_h, $\mathbf{u}_h$, and λ_H are locally approximated by polynomials of degree $\ell-1$, $k-1$, and $m-1$, respectively, then there exists C, independent of h and H, such that*

$$\begin{aligned}
&\|\nabla\cdot(\mathbf{u}-\mathbf{u}_h)\|_0 \le C\|f\|_\ell h^\ell,\\
&\|\mathbf{u}-\mathbf{u}_h\|_0 \le C\big\{\|\mathbf{u}\|_k h^k + \|p\|_{m+1/2}H^{m-1/2} + \|\mathbf{u}\|_{k+1/2}h^k H^{1/2}\big\},\\
&\|p-p_h\|_0 \le C\big\{\|p\|_\ell h^\ell + \|p\|_{m+1/2}H^{m+1/2}\\
&\qquad\qquad + (\|f\|_\ell h^\ell + \|\mathbf{u}\|_k h^k)H + \|\mathbf{u}\|_{k+1/2}h^k H^{3/2}\big\},
\end{aligned}$$

where the last estimate requires that Ω support elliptic regularity.

The velocity estimate is formally $\mathscr{O}(h^k + H^{m-1/2})$, but of course it suffers from the problem of scale when $h < \varepsilon < H$. Since λ is defined on Γ, we lose $1/2$-derivative going to Ω, or $H^{-1/2}$ in the estimates. Thus when h and H are of the same order, we need polynomials of degree 1 (actually 1/2) more for the mortar approximation [9].

The error estimate bounds given in Theorem 13 depend on the solution, and so may be very large for our heterogeneous problem. To deal with this problem of scale, one can use a-posteriori analysis to adapt the meshes to the scales of the system [9, 58]. One can also define a mortar space based on homogenization, which results in an error estimate for the two-scale separation case that is optimal and has no numerical resonance term [12]. From Theorem 2,

$$\lambda \approx p_\varepsilon^1 = \Big(1+\varepsilon\sum_{j=1}^{d}\omega_j(x,x/\varepsilon)\,\frac{\partial}{\partial x_j}\Big)p_0(x), \tag{77}$$

so we define M_H by replacing p_0 above by piecewise polynomials and restrict back to Γ.

8 Some Numerical Results

We present two numerical test examples that model incompressible single phase flow in a porous medium. The permeability fields are geostatistically generated, with the first being statistically homogeneous. Since homogenization theory extends to this case, the multiscale convergence results relating the coarse grid spacing h to ε apply, as do the techniques mentioned that are defined using homogenization theory. The other test case has a permeability that is highly nonhomogeneous in the statistical sense, and so the convergence results do not strictly apply. That is, we are far from the case of scale separation. Nevertheless, the example is presented to demonstrate that the methods presented herein can work well even in these cases.

In each example, the domain is a rectangle in $\mathbb{R}^2$. The permeability is defined as a piecewise constant on a fine uniform rectangular grid. Moreover, the source function f is positive in the lower left corner element and of equal strength but negative in the upper right corner element.

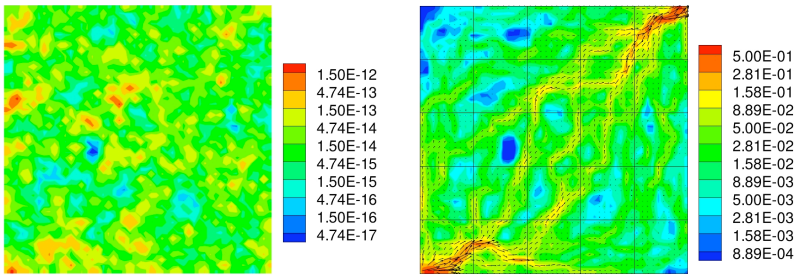
We solve each problem nine times, divided into three sets of experiments. For the first set of experiments, we solve the problem on the fine grid using standard RT0 and BDM1 mixed elements. This gives us the "true" or reference solution that the multiscale techniques should approximate. We take the RT0 results as the reference solution.

For the second set of experiments, we solve the problem using the variational multiscale method on a coarsened grid that leaves a 10×10 subgrid; that is, we use a factor of 100 upscaling. We use the multiscale finite elements defined in Sect. 5, in particular, the standard ME0 and ME1 multiscale finite elements, as well as the multiscale dual support elements MD and the ones defined through homogenization theory HE.

For the final set of experiments, we solve each problem using the domain decomposition mortar method, using subdomains with the same 10×10 subgrid. We use a single mortar element along the subdomain (or coarse grid) edges, with the mortar space defined as the piecewise discontinuous linear or quadratic functions P1M and P2M, as well as the mortar space defined using homogenization HM (77).

The upscaling operator or subgrid problems are solved on the fine 10×10 grid using RT0.

8.1 Example 1: A Statistically Homogeneous Permeability



Permeability a_ε BDM1 velocity $\mathbf{u}$

Fig. 10 Ex. 1. Left: The permeability on a log scale. Right: The velocity $\mathbf{u}$ as computed on the fine scale using BDM1. The arrows show the velocity and the color shows its magnitude (speed) on a log scale.

In the first example, the permeability field is a scalar field with a variation of about 5 orders of magnitude. It is depicted in Fig. 10 on a base-10 log scale. The fine grid is 50×50, and the coarse grid is only 5×5. We also show the fine-scale

velocity using BDM1. Note that the color is the speed $|\mathbf{u}|$, on a log scale, while the arrows, barely visible, show the velocity itself. Thus nearly all the flow concentrates in channels, depicted in warmer colors, from the lower left injection well to the upper right extraction well.

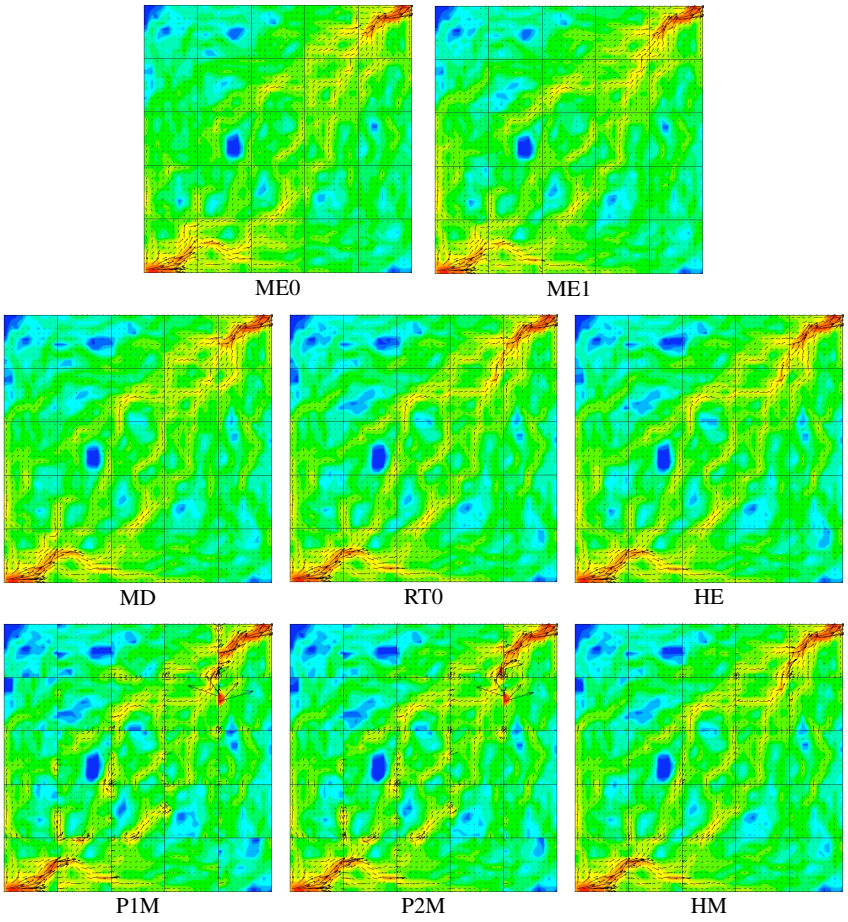


Fig. 11 Ex. 1. The arrows show the velocity $\mathbf{u}$, and the color depicts its magnitude (speed) on a log scale. The center plot shows the fine-scale RT0 velocity, with the multiscale finite elements beside and above it. The mortar methods are on the bottom row.

In Fig. 11 we show results for the seven multiscale methods, and the fine-scale RT0 result in the center plot (which is nearly identical to the BDM1 result). The variational multiscale method using various multiscale elements have been plotted beside and above the RT0 result for easy comparison. The mortar results appear in a row below RT0. All the methods appear to do well to the eye.

Table 1 Ex. 1. Relative errors with respect to the fine-scale RT0 solution for pressure p and velocity $\mathbf{u}$, measured in both L^2 and L^∞ norms for three sets of methods. First, BDM1 is obtained on the fine grid, and differs by a small amount from RT0. Next are the multiscale finite element methods, and the number of coarse degrees of freedom (DOF) used by each method is noted. Finally, the mortar methods are shown with their number of DOF per mortar interface (i.e., coarse grid edge).

Method	DOF per coarse edge	Pressure error L^2	Pressure error L^∞	Velocity error L^2	Velocity error L^∞
BDM1	—	0.043	0.041	0.036	0.028
ME0	1	0.329	0.329	0.300	0.343
ME1	2	0.150	0.144	0.249	0.345
MD	1	0.068	0.105	0.170	0.136
HE	2	0.054	0.084	0.088	0.066
P1M	2	0.157	0.139	0.388	0.948
P2M	3	0.099	0.097	0.323	0.819
HM	3	0.006	0.012	0.041	0.044

Considering the fine-scale RT0 results as the reference solution, we give the relative errors of the other eight methods in Table 1. The methods differ in the number of degrees of freedom each has on the coarse interfaces, so this is given in one column of the table. Next we show errors in both pressure and velocity, though velocity is normally the more important quantity. The errors are measured in the L^2- and L^∞-norms. Note that RT0 and BDM1 disagree in velocity by 3.6% in L^2, indicating that the problem is difficult to resolve on our 50×50 mesh, and giving an idea of the error that we should tolerate in these applications.

The basic methods ME0, ME1, P1M, and P2M are actually relatively poor, giving 25% to 39% relative L^2 velocity errors. If one looks carefully at Fig. 11, one can see the differences between these methods and RT0. ME0 and ME1 are too numerically diffuse, and P1M and P2M mainly have difficulties in a few isolated points in the domain. The remaining three methods are quite reasonable, both in terms of the numerical error and in the velocity plots. The MD method has 17% L^2 velocity error, followed by the HE method with 9% error, and finally the HM mortar method does the best job at 4% relative velocity error.

8.2 Example 2: A Statistically Nonhomogeneous Permeability

The second numerical test example is based on the Tenth Society of Petroleum Engineers Comparative Solution Project [30]. The project includes a difficult three-dimensional permeability field. We take one 60×220 two-dimensional slice, the twentieth, which represents a near shore environment with definite local channeling. It is badly nonisotropic, as depicted in Fig. 12. The permeability is a diagonal tensor, so we show the two components of permeability, which vary on a log scale by about 5 orders of magnitude. As one can see from the BDM1 fine-scale solution (we show only the speed $|\mathbf{u}|$, again on a log scale), the flow field is quite complex. Generally speaking, fluid that enters the domain at the lower left corner injection

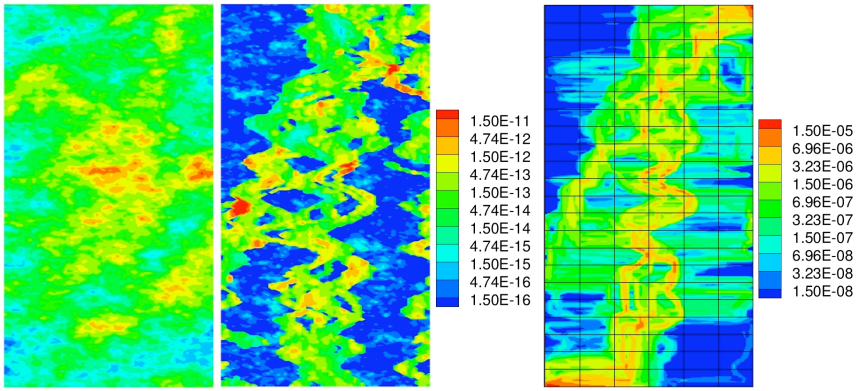


Fig. 12 Ex. 2. Left: The anisotropic x- and y-permeabilities on a log scale. Right: The speed $|\mathbf{u}|$ as computed on the fine scale using BDM1 on a log scale.

well cannot travel upwards in the y-direction until it travels in x a good third of the domain. It then experiences a greater y permeability, and so flows along the permeability channels to the upper right corner extraction well.

The other eight techniques' speeds are shown in Fig. 13, and the relative errors are given in Table 2. As we surmised from the permeability field, we have now quantified that this is a very difficult problem: the BDM1 solution has 8.6% relative velocity error in L^2 compared to RT0, even though each is solved on the fine grid.

Table 2 Ex. 2. Relative errors with respect to the fine-scale RT0 solution for pressure p and velocity $\mathbf{u}$, measured in both L^2 and L^∞ norms for three sets of methods. First, BDM1 is obtained on the fine grid, and differs somewhat from RT0. Next are the multiscale finite element methods, and the number of coarse degrees of freedom (DOF) used by each method is noted. Finally, the mortar methods are shown with their number of DOF per mortar interface (i.e., coarse grid edge).

Method	DOF per coarse edge	Pressure error L^2	Pressure error L^∞	Velocity error L^2	Velocity error L^∞
BDM1	—	0.047	0.052	0.086	0.164
ME0	1	2.454	0.414	0.717	0.598
ME1	2	0.799	0.194	0.625	0.443
MD	1	0.223	0.332	0.453	0.575
HE	2	0.149	0.197	0.303	0.261
P1M	2	0.132	0.056	0.418	0.748
P2M	3	0.075	0.038	0.306	0.441
HM	3	0.027	0.023	0.142	0.198

The multiscale finite elements ME0 and ME1 have very large 72% and 63% errors, respectively. The figure shows that they give speeds that are much too diffuse compared to RT0.

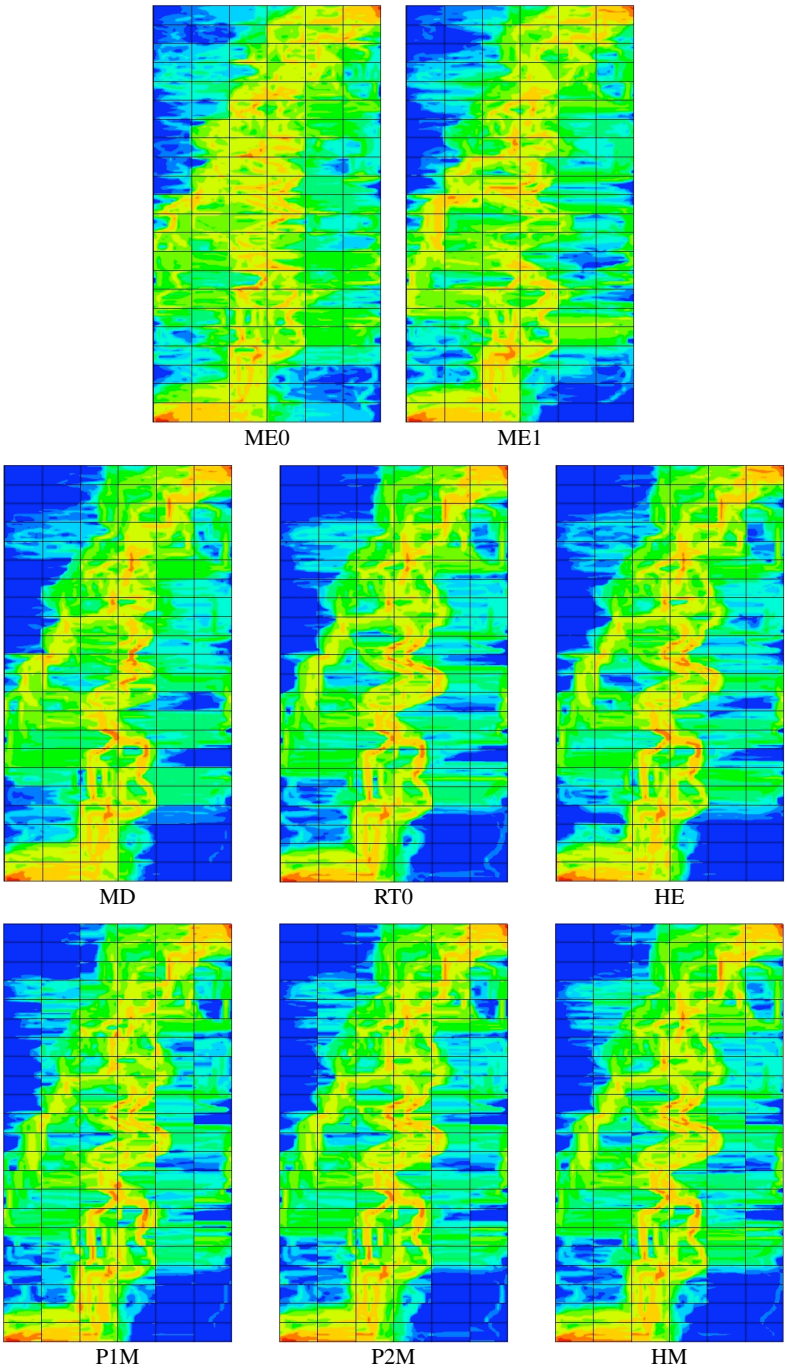


Fig. 13 Ex. 2. The speed $|\mathbf{u}|$ on a log scale. The center plot shows the fine-scale RT0 velocity, with the multiscale finite elements beside and above it. The mortar methods are on the bottom row.

The multiscale elements MD and HE do much better (45% and 30% error, respectively). The mortar methods P1M and P2M are quite similar (42% and 31% error, respectively). Each of these four methods have speed plots that match RT0 quite well to the eye, and in particular, match the high-speed channel flow quite well. These solutions would be sufficient for preliminary engineering analyses and some stochastic simulation studies. However, only the mortar method based on homogenization theory [12] has a reasonable error at 14%. Perhaps to the eye it is also the closest match to RT0.

8.3 Some Techniques for Controlling Errors

Because multiscale methods are reduced degree of freedom methods, they are subject to error in somewhat unforeseen ways. A clear example was seen in Fig. 11, where both P1M and P2M have trouble at a spot on the coarse interface parallel to y between subdomains $(4,4)$ and $(5,4)$, counting from the lower left. It is not clear why this spot causes trouble for these two methods but not the others.

There has been a great deal of recent work on trying to mitigate this problem. Several authors advocate the use of limited global information to improve the definition of multiscale finite elements [3, 27, 34, 2]. The method seems most useful in the case of nonlinear problems, in which case one solves a global fine-scale *linear* problem, and time dependent and stochastic problems, in which case one solves only one or a few global fine-scale problems. The global information can be used to better define the multiscale finite element on the boundaries of the coarse elements, and then one uses energy minimizing extension into the interior. Some works deal with adaptive methods and a-posteriori error estimation and control of errors [11, 52, 57, 58]. Basically, one includes more scales where the errors are estimated to be large. Other works deal with using the ideas of multiscale finite elements and domain decomposition to define better iterative solvers for the fine-scale system [55, 65, 54, 59, 60, 41, 53]. The idea is to iterate on the fine-scale system to convergence using multiscale ideas basically as a preconditioner or in defining prolongation and restriction operators in multigrid.

Acknowledgements This work was supported by the U.S. National Science Foundation and the Center for Frontiers of Subsurface Energy Security, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001114.

References

1. J. E. Aarnes and B.-O. Heimsund. Multiscale discontinuous Galerkin methods for elliptic problems with multiple scales. In Timothy J. Barth et al., editors, *Multiscale Methods in*

Science and Engineering, volume 44 of *Lecture Notes in Computational Science and Engineering*, pages 1–20. Springer Berlin Heidelberg, 2005.
2. J. E. Aarnes, Y. Efendiev, and L. Jiang. Mixed multiscale finite element methods using limited global information. *Multiscale Model. Simul.*, 7(2):655–676, 2008.
3. J. E. Aarnes. On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation. *Multiscale Model. Simul.*, 2(3):421–439, 2004.
4. J. E. Aarnes, S. Krogstad, and K.-A. Lie. A hierarchical multiscale method for two-phase flow based upon mixed finite elements and nonuniform coarse grids. *Multiscale Model. Simul.*, 5:337–363, 2006.
5. T. Arbogast. Numerical subgrid upscaling of two-phase flow in porous media. In Z. Chen, R. E. Ewing, and Z.-C. Shi, editors, *Numerical treatment of multiphase flows in porous media*, volume 552 of *Lecture Notes in Physics*, pages 35–49. Springer, Berlin, 2000.
6. T. Arbogast. Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems. *SIAM J. Numer. Anal.*, 42:576–598, 2004.
7. T. Arbogast. Homogenization-based mixed multiscale finite elements for problems with anisotropy. *Submitted*, 2010.
8. T. Arbogast and K. J. Boyd. Subgrid upscaling and mixed multiscale finite elements. *SIAM J. Numer. Anal.*, 44(3):1150–1171, 2006.
9. T. Arbogast, L. C. Cowsar, M. F. Wheeler, and I. Yotov. Mixed finite element methods on non-matching multiblock grids. *SIAM J. Numer. Anal.*, 37:1295–1315, 2000.
10. T. Arbogast, S. E. Minkoff, and P. T. Keenan. An operator-based approach to upscaling the pressure equation. In V. N. Burganos et al., editors, *Computational Methods in Water Resources XII, Vol. 1: Computational Methods in Contamination and Remediation of Water Resources*, pages 405–412, Southampton, U.K., 1998. Computational Mechanics Publications.
11. T. Arbogast, G. Pencheva, M. F. Wheeler, and I. Yotov. A multiscale mortar mixed finite element method. *Multiscale Model. Simul.*, 6(1):319–346, 2007.
12. T. Arbogast and Hailong Xiao. A multiscale mortar mixed space based on homogenization for heterogeneous elliptic problems. *In preparation*, 2010.
13. I. Babuška. The finite element method with Lagrangian multipliers. *Numer. Math.*, 20:179–192, 1973.
14. I. Babuška, G. Caloz, and J. E. Osborn. Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.*, 31:945–981, 1994.
15. I. Babuška and R. Lipton. Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. Technical Report 10–12, Institute for Computational Engineering and Sciences, Univ. of Texas, Austin, Texas, USA, Mar. 2010.
16. I. Babuška and J. E. Osborn. Generalized finite element methods: their performance and their relation to mixed methods. *SIAM J. Numer. Anal.*, 20:510–536, 1983.
17. J. Bear. *Dynamics of Fluids in Porous Media*. Dover, New York, 1972.
18. J. Bear and A. H.-D. Cheng. *Modeling Groundwater Flow and Contaminant Transport*. Springer, New York, 2010.
19. A. Bensoussan, J. L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structure*. North-Holland, Amsterdam, 1978.
20. C. Bernardi, Y. Maday, and A. T. Patera. A new nonconforming approach to domain decomposition: The mortar element method. In H. Brezis and J. L. Lions, editors, *Nonlinear partial differential equations and their applications*. Longman Scientific & Technical, UK, 1994.
21. S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York, 1994.
22. F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *RAIRO*, 8:129–151, 1974.
23. F. Brezzi. Interacting with the subgrid world. In *Numerical Analysis, 1999*, pages 69–82. Chapman and Hall, 2000.
24. F. Brezzi, J. Douglas, Jr., R. Duràn, and M. Fortin. Mixed finite elements for second order elliptic problems in three variables. *Numer. Math.*, 51:237–250, 1987.

25. F. Brezzi, J. Douglas, Jr., and L. D. Marini. Two families of mixed elements for second order elliptic problems. *Numer. Math.*, 47:217–235, 1985.
26. F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Springer-Verlag, New York, 1991.
27. Y. Chen and L. J. Durlofsky. Adaptive local-global upscaling for general flow scenarios in heterogeneous formations. *Transp. Por. Med.*, 62:157–185, 2006.
28. Z. Chen and T. Y. Hou. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Math. Comp.*, 72:541–576, 2003.
29. Z. Chen, G. Huan, and Y. Ma. *Computational Methods for Multiphase Flows in Porous Media*, volume 2 of *Computational Science and Engineering Series*. SIAM, Philadelphia, 2006.
30. M. A. Christie. Tenth spe comparative solution project: A comparison of upscaling techniques, SPE 72469. *SPE Res. Eval. & Engng.*, 2001. 2001 SPE Reservoir Simulation Symposium.
31. Ph. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.
32. Weinan E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1:87–132, 2003.
33. Y. Efendiev, J. Galvis, and Xiao-Hui Wu. Multiscale finite element methods for high-contrast problems using local spectral basis functions. *J. Comput. Phys.*, 230(4):937–955, 2011.
34. Y. Efendiev, V. Ginting, T. Y. Hou, and R. E. Ewing. Accurate multiscale finite element methods for two-phase flow simulations. *J. Comput. Phys.*, 220(1):155–174, 2006.
35. Y. R. Efendiev, T. Y. Hou, and X.-H. Wu. Convergence of a nonconforming multiscale finite element method. *SIAM J. Numer. Anal.*, 37:888–910, 2000.
36. G. B. Folland. *Introduction to Partial Differential Equations*. Princeton, 1976.
37. B. Ganis and I. Yotov. Implementation of a mortar mixed finite element method using a multiscale flux basis. *Comput. Methods Appl. Mech. Engrg.*, 198:3989–3998, 2009.
38. D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, Berlin, 1983.
39. V. Ginting. Analysis of two-scale finite volume element method for elliptic problem. *J. Numer. Math.*, 12(2):119–141, 2004.
40. R. Glowinski and M. F. Wheeler. Domain decomposition and mixed finite element methods for elliptic problems. In R. Glowinski et al., editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 144–172. SIAM, Philadelphia, 1988.
41. I. G. Graham and R. Scheichl. Robust domain decomposition algorithms for multiscale PDEs. *Numer. Meth. Partial Diff. Eqns.*, 23(4):859–878, 2007.
42. P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston, 1985.
43. M. A. Hesse, B. T. Mallison, and H. A. Tchelepi. Compact multiscale finite volume method for heterogeneous anisotropic elliptic equations. *Multiscale Model. Simul.*, 7(2):934–962, 2008.
44. U. L. Hetmaniuk and R. B. Lehoucq. A special finite element methods based on component mode synthesis techniques. *ESAIM: Math. Modelling and Numer. Anal.*, 2010.
45. U. Hornung, editor. *Homogenization and Porous Media*. Interdisciplinary Applied Mathematics Series. Springer-Verlag, New York, 1997.
46. T. Y. Hou and X. H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134:169–189, 1997.
47. T. Y. Hou, X.-H. Wu, and Z. Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comp.*, 68:913–943, 1999.
48. 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.*, 127:387–401, 1995.
49. 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.*, 166:3–24, 1998.
50. P. Jenny, S. H. Lee, and H. A. Tchelepi. Multi-scale finite-volume method for elliptic problems in subsurface flow simulation. *J. Comp. Phys.*, 187:47–67, 2003.

51. V. V. Jikov, S. M. Kozlov, and O. A. Oleinik. *Homogenization of Differential Operators and Integral Functions*. Springer-Verlag, New York, 1994.
52. M. G. Larson and A. Målqvis. Adaptive variational multiscale methods based on a posteriori error estimation: energy norm estimates for elliptic problems. *Comput. Methods Appl. Mech. Engrg.*, 196(21–24):2313–2324, 2007.
53. J. Van Lent, R. Scheichl, and I. G. Graham. Energy minimizing coarse spaces for two-level Schwarz methods for multiscale PDEs. *Numer. Lin. Alg. with Applic.*, 16(10):775–799, 2009.
54. S. P. MacLachlan and J. D. Moulton. Multilevel upscaling through variational coarsening. *Water Resour. Res.*, 42, 2006.
55. J. D. Moulton, Jr. J. E. Dendy, and J. M. Hyman. The black box multigrid numerical homogenization algorithm. *J. Comput. Phys.*, 142(1):80–108, 1998.
56. J. Nolen, G. Papanicolaou, and O. Pironneau. A framework for adaptive multiscale methods for elliptic problems. *Multiscale Model. Simul.*, 7(1):171–196, 2008.
57. J. M. Nordbotten. Adaptive variational multiscale methods for multiphase flow in porous media. *Multiscale Model. Simul.*, 7(3):1455–1473, 2009.
58. G. Pencheva, M. Vohralik, M. F. Wheeler, and T. Wildey. Robust a posteriori error control and adaptivity for multiscale, multinumerics, and mortar coupling. *Submitted*, 2010.
59. J. M. Rath. Darcy flow, multigrid, and upscaling. In et al. W. W. Hager, editor, *Multiscale Optimization Methods and Applications*, volume 82 of *Nonconvex Optimization and its Applications*, pages 337–366. Springer, New York, 2006.
60. J. M. Rath. *Multiscale Basis Optimization for Darcy Flow*. PhD thesis, Univ. of Texas, Austin, Texas, May 2007.
61. R. A. Raviart and J. M. Thomas. A mixed finite element method for 2nd order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical Aspects of Finite Element Methods*, number 606 in Lecture Notes in Math., pages 292–315. Springer-Verlag, New York, 1977.
62. E. Sanchez-Palencia. *Non-homogeneous Media and Vibration Theory*. Number 127 in Lecture Notes in Physics. Springer-Verlag, New York, 1980.
63. H. A. Schwarz. Gesammelte mathematische adhandlungen. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 15:272–286, 1870.
64. T. Strouboulis, K. Copps, and I. Babuška. The generalized finite element method. *Comput. Methods Appl. Mech. Engrg.*, 190:4081–4193, 2001.
65. Jinchao Xu and L. Zikatanov. On an energy minimizing basis for algebraic multigrid methods. *Comput. Vis. Sci.*, 7(3–4):121–127, 2004.