

MULTISCALE FINITE ELEMENT METHODS FOR NONLINEAR PROBLEMS AND THEIR APPLICATIONS *

Y. EFENDIEV [†], T. HOU [‡], AND V. GINTING [§]

Abstract. In this paper we propose a generalization of multiscale finite element methods (MsFEM) to nonlinear problems. We study the convergence of the proposed method for nonlinear elliptic equations and propose an oversampling technique. Numerical examples demonstrate that the oversampling technique greatly reduces the error. The application of MsFEM to porous media flows is considered. Finally, we describe further generalizations of MsFEM to nonlinear time-dependent equations and discuss the convergence of the method for various kinds of heterogeneities.

Key words. multiscale, finite element, upscaling, nonlinear, elliptic, oversampling

AMS Classification number: 65N99

1. Introduction

Many processes involve a wide range of scales. Because of the scale disparity in multiscale problems, it is often impossible to resolve the effects of small scales directly. For this reason some type of coarsening or upscaling is performed. The main idea of upscaling techniques is to form coarse-scale equations with a prescribed analytical form that may differ from the underlying fine-scale equations. In multiscale methods, by contrast, the fine-scale information may be carried throughout the simulation, and the coarse-scale equations are generally not expressed analytically, but rather formed and solved numerically.

Recently, a number of multiscale numerical methods, such as residual free bubbles [5, 31], variational multiscale method [21], multiscale finite element method (MsFEM) [19], two-scale finite element methods [25], two-scale conservative subgrid approaches [1, 2], and heterogeneous multiscale method (HMM) [10] have been proposed. We remark that special base functions in finite element methods have been used earlier in [4] and in [3], where using a special base function, the generalized finite element methods is introduced. In this paper we will generalize MsFEM to nonlinear problems. Originally, MsFEM is proposed for linear equations and its main idea is to use oscillatory base functions to capture the local-scale information. The pre-computed multiscale base functions allow us to interpolate a coarse-scale function, defined at the nodal values of the coarse grid, to the underlying fine grid. This idea can be naturally generalized to nonlinear problems if one considers, instead of the base functions, a multiscale map from the coarse grid space to the underlying fine grid space. This multiscale map is constructed using the solutions of the local problems and provides us with the interpolation of the coarse-scale function, defined at the nodal values of the coarse grid, to the underlying fine grid. For linear problems, our multiscale map is linear and, thus, the image of the coarse dimensional space is a linear space with the same dimension. A basis for the multiscale space can be obtained by mapping a basis of the coarse dimensional space. The latter gives us the multiscale finite element basis functions introduced in [19]. Once the multiscale mapping is defined, we formulate the global finite element formulation of the problem. Our multiscale finite element

*Received: June 22, 2004; accepted (in revised version): September 28, 2004. Communicated by Shi Jin.

[†]Department of Mathematics, Texas A&M University, College Station, TX 77843-3368.

[‡]Applied Mathematics, California Institute of Technology, Pasadena, CA 91125.

[§]Department of Mathematics, Texas A&M University, College Station, TX 77843-3368.

methods use a Petrov-Galerkin formulation in which we use multiscale finite element bases as basis functions and standard linear finite elements as test functions. We note that the Petrov-Galerkin formulation of MsFEM is found to have an advantage [18] for linear problems. We would like to stress that the formulation of MsFEM does not require any assumptions on the nature of heterogeneities, such as periodicity, almost periodicity, etc.

We consider the analysis of MsFEM for general nonlinear elliptic equations, $u_\epsilon \in W_0^{1,p}(\Omega)$

$$-\operatorname{div}(a_\epsilon(x, u_\epsilon, D_x u_\epsilon)) + a_{0,\epsilon}(x, u_\epsilon, D_x u_\epsilon) = f, \quad (1.1)$$

where $a_\epsilon(x, \eta, \xi)$ and $a_{0,\epsilon}(x, \eta, \xi)$, $\eta \in \mathbb{R}$, $\xi \in \mathbb{R}^d$ satisfy some assumptions given by (3.1)-(3.5), which guarantee the well-posedness of the nonlinear elliptic problem (1.1). Here $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain and ϵ denotes the small scale of the problem. The homogenization of nonlinear partial differential equations has been studied previously (see, e.g., [28]). It can be shown that a solution u_ϵ converges (up to a sub-sequence) to u in an appropriate norm, where $u \in W_0^{1,p}(\Omega)$ is a solution of a homogenized equation

$$-\operatorname{div}(a^*(x, u, Du)) + a_0^*(x, u, Du) = f. \quad (1.2)$$

The homogenized coefficients can be computed if we make an additional assumption on the heterogeneities, such as periodicity, almost periodicity, or when the fluxes are strictly stationary fields with respect to spatial variables. In these cases, an auxiliary problem is formulated and used in the calculations of the homogenized fluxes, a^* and a_0^* . Our motivation in considering this type of equation stems from porous media applications, where nonlinear fluxes arise. In particular, we are interested in porous media flows in unsaturated media and the transport of two-phase flows in heterogeneous porous media. In these examples, nonlinearities arise due to the interaction between the phases and components.

In this paper we study the convergence of the generalized MsFEM for periodic heterogeneities. To analyze the method, we first approximate the solutions of the local problems by introducing appropriate correctors, which are periodic with respect to the fast variables. These approximations of the local solutions allow us to extract the homogenized behavior of MsFEM solutions and compare it with the homogenized solutions of the continuous equations. Sharp estimates for the corrector approximations are obtained in the paper. The analysis allows us to understand the resonance error and propose an oversampling technique as in [19]. Numerical examples are presented in the paper to show the accuracy of the oversampling method. We use both periodic and random fields with long-range correlation structures (with and without discontinuities) in our numerical experiments. We present numerical examples for both multiscale finite element and multiscale finite volume element methods. Multiscale finite volume element methods are very closely related to multiscale finite element method, where the formulation of the method follows the standard finite volume element methods. All the examples clearly demonstrate the advantages of the oversampling method. In particular, the oversampling approach provides small errors for relatively large coarsening. Further generalization of the analysis to the cases of more general heterogeneities is discussed in the paper. Finally, we would like to note that the resonance errors are a common feature of multiscale methods unless periodic problems are considered and the solutions of the local problems in an exact period are used. In this case, one can solve the local problems in one period to approximate the multiscale map.

The paper is organized in the following way. In the next section, we introduce MsFEM for nonlinear problems. Section 3 is devoted to the analysis of MsFEM. In Section 4, numerical examples are presented. In particular, we show that with the oversampling technique, the error is reduced dramatically. The applications of MsFEM to porous media flows are also considered in Section 4. Finally, in Section 5 some conclusions are drawn. We present further generalizations of MsFEM to nonlinear parabolic equations and discuss the convergence of the method for various types of heterogeneities.

2. Multiscale finite element methods (MsFEM)

The goal of MsFEM is to find a numerical approximation of a homogenized solution *without* solving auxiliary problems (e.g., periodic cell problems) that arise in homogenization. The homogenized solutions are sought on a coarse grid space S^h , where $h \gg \epsilon$. Let \mathbf{K}^h be a partition of Ω . We denote by S^h standard family of finite dimensional space, which possesses approximation properties, e.g., piecewise linear functions over triangular elements,

$$S^h = \{v_h \in C^0(\bar{\Omega}) : \text{the restriction } v_h \text{ is linear for each element } K \text{ and } v_h = 0 \text{ on } \partial\Omega\}. \tag{2.1}$$

In further presentation, K is a triangular element that belongs to \mathbf{K}^h . To formulate MsFEM for general nonlinear problems, we will need (1) a *multiscale mapping* that gives us the desired approximation containing the small scale information and (2) a *multiscale numerical formulation* of the equation.

Multiscale mapping. Introduce the mapping $E^{MsFEM} : S^h \rightarrow V_\epsilon^h$ in the following way. For each element $v_h \in S^h$, $v_{\epsilon,h} = E^{MsFEM} v_h$ is defined as the solution of

$$-\text{div}(a_\epsilon(x, \eta^{v_h}, D_x v_{\epsilon,h})) = 0 \text{ in } K, \tag{2.2}$$

$v_{\epsilon,h} = v_h$ on ∂K and $\eta^{v_h} = \frac{1}{|K|} \int_K v_h dx$ for each K . We would like to point out that different boundary conditions can be chosen to obtain more accurate solutions and this will be discussed later. Note that for linear problems, E^{MsFEM} is a linear operator, where for each $v_h \in S^h$, $v_{\epsilon,h}$ is the solution of the linear problem. Consequently, V_ϵ^h is a linear space that can be obtained by mapping a basis of S^h . This is precisely the construction presented in [19] for linear elliptic equations.

Multiscale numerical formulation. Multiscale finite element formulation of the problem is the following. Find $u_h \in S^h$ (consequently, $u_{\epsilon,h} (= E^{MsFEM} u_h) \in V_\epsilon^h$) such that

$$\langle A_{\epsilon,h} u_h, v_h \rangle = \int_\Omega f v_h dx \quad \forall v_h \in S^h, \tag{2.3}$$

where

$$\langle A_{\epsilon,h} u_h, v_h \rangle = \sum_{K \in \mathbf{K}^h} \int_K ((a_\epsilon(x, \eta^{u_h}, D_x u_{\epsilon,h}), D_x v_h) + a_{0,\epsilon}(x, \eta^{u_h}, D_x u_{\epsilon,h}) v_h) dx. \tag{2.4}$$

Note that the above formulation of MsFEM is a generalization of the Petrov-Galerkin MsFEM introduced in [18] for linear problems. MsFEM, introduced above, can be generalized to different kinds of nonlinear problems and this will be discussed later.

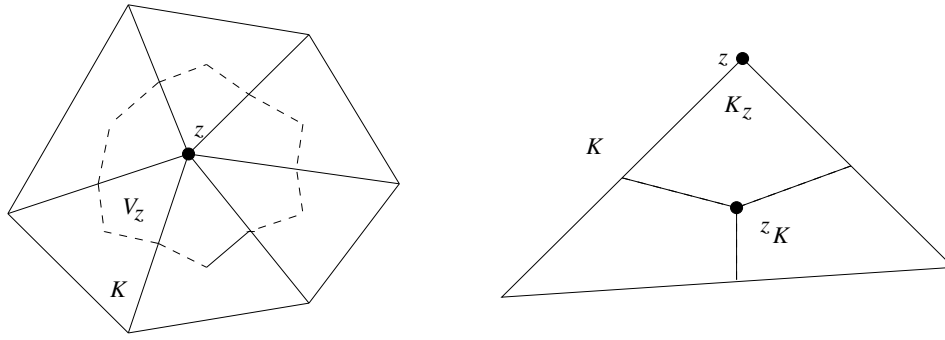


FIG. 2.1. Left: Portion of triangulation sharing a common vertex z and its control volume. Right: Partition of a triangle K into three quadrilaterals

2.1. Multiscale finite volume element method (MsFVEM). The formulation of multiscale finite element (MsFEM) can be extended to a finite volume method. By its construction, the finite volume method has local conservative properties [16] and it is derived from a local relation, namely the balance equation/conservation expression on a number of subdomains which are called control volumes. Finite volume element method can be considered as a Petrov-Galerkin finite element method, where the test functions are constants defined in a dual grid. Consider a triangle K , and let z_K be its barycenter. The triangle K is divided into three quadrilaterals of equal area by connecting z_K to the midpoints of its three edges. We denote these quadrilaterals by K_z , where $z \in Z_h(K)$ are the vertices of K . Also we denote $Z_h = \bigcup_K Z_h(K)$, and Z_h^0 are all vertices that do not lie on Γ_D , where Γ_D is Dirichlet boundaries. The control volume V_z is defined as the union of the quadrilaterals K_z sharing the vertex z (see Figure 2.1). The multiscale finite volume element method (MsFVEM) is to find $u_h \in S^h$ (consequently, $u_{\epsilon,h} = E^{MsFVEM} u_h$ such that

$$-\int_{\partial V_z} a_\epsilon(x, \eta^{u_h}, D_x u_{\epsilon,h}) \cdot n dS + \int_{V_z} a_{0,\epsilon}(x, \eta^{u_h}, D_x u_{\epsilon,h}) dx = \int_{V_z} f dx \quad \forall z \in Z_h^0, \tag{2.5}$$

where n is the unit normal vector pointing outward on ∂V_z . Note that the number of control volumes that satisfy (2.5) is the same as the dimension of S^h . We will present numerical results for both multiscale finite element and multiscale finite volume element methods.

2.2. Examples of V_ϵ^h . Linear case. For linear operators, V_ϵ^h can be obtained by mapping a basis of S^h . Define a basis of S^h , $S^h = \text{span}(\phi_0^i)$, where ϕ_0^i are standard linear basis functions. In each element $K \in \mathbf{K}^h$, we define a set of nodal basis $\{\phi_\epsilon^i\}$, $i = 1, \dots, n_d$ with $n_d (= 3)$ being the number of nodes of the element, satisfying

$$-\text{div}(a_\epsilon(x) D_x \phi_\epsilon^i) = 0 \quad \text{in } K \in \mathbf{K}^h \tag{2.6}$$

and $\phi_\epsilon^i = \phi_0^i$ on ∂K . Thus, we have

$$V_\epsilon^h = \text{span}\{\phi_\epsilon^i; i = 1, \dots, n_d, K \in \mathbf{K}^h\} \subset H_0^1(\Omega).$$

Oversampling technique can be used to improve the method [19].

Special nonlinear case. For the special case, $a_\epsilon(x, u_\epsilon, D_x u_\epsilon) = a_\epsilon(x)b(u_\epsilon)D_x u_\epsilon$, V_ϵ^h can be related to the linear case. Indeed, for this case, the local problems associated with the multiscale mapping E^{MsFEM} (see (2.2)) have the form

$$-\operatorname{div}(a_\epsilon(x)b(\eta^{v_h})D_x v_{\epsilon,h}) = 0 \text{ in } K.$$

Because η^{v_h} are constants over K , the local problems satisfy the linear equations,

$$-\operatorname{div}(a_\epsilon(x)D_x \phi_\epsilon^i) = 0 \text{ in } K,$$

and V_ϵ^h can be obtained by mapping a basis of S^h as it is done for the first example. Thus, for this case one can construct the base functions in the beginning of the computations.

V_ϵ^h using subdomain problems. One can use the solutions of smaller (than $K \in \mathbf{K}^h$) subdomain problems to approximate the solutions of the local problems (2.2). This can be done in various ways based on a homogenization expansion. For example, instead of solving (2.2) we can solve (2.2) in a subdomain S with boundary conditions v_h restricted onto the subdomain boundaries, ∂S . Then the gradient of the solution in a subdomain can be extended periodically to K to approximate $D_x v_{\epsilon,h}$ in (2.4). $v_{\epsilon,h}$ can be easily reconstructed based on $D_x v_{\epsilon,h}$. When the multiscale coefficient has a periodic structure, the multiscale mapping can be constructed over one periodic cell with a specified average. In this case, $v_{\epsilon,h}$ is approximated by \mathcal{P} , which is defined by (3.18).

3. Analysis of MsFEM

For the analysis of MsFEM, we assume the following conditions for $a_\epsilon(x, \eta, \xi)$ and $a_{0,\epsilon}(x, \eta, \xi)$, $\eta \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$.

$$|a_\epsilon(x, \eta, \xi)| + |a_{0,\epsilon}(x, \eta, \xi)| \leq C(1 + |\eta|^{p-1} + |\xi|^{p-1}), \tag{3.1}$$

$$(a_\epsilon(x, \eta, \xi_1) - a_\epsilon(x, \eta, \xi_2), \xi_1 - \xi_2) \geq C|\xi_1 - \xi_2|^p, \tag{3.2}$$

$$(a_\epsilon(x, \eta, \xi), \xi) + a_{0,\epsilon}(x, \eta, \xi)\eta \geq C|\xi|^p. \tag{3.3}$$

Denote

$$H(\eta_1, \xi_1, \eta_2, \xi_2, r) = (1 + |\eta_1|^r + |\eta_2|^r + |\xi_1|^r + |\xi_2|^r), \tag{3.4}$$

for arbitrary $\eta_1, \eta_2 \in \mathbb{R}$, $\xi_1, \xi_2 \in \mathbb{R}^d$, and $r > 0$. We further assume that

$$\begin{aligned} & |a_\epsilon(x, \eta_1, \xi_1) - a_\epsilon(x, \eta_2, \xi_2)| + |a_{0,\epsilon}(x, \eta_1, \xi_1) - a_{0,\epsilon}(x, \eta_2, \xi_2)| \\ & \leq CH(\eta_1, \xi_1, \eta_2, \xi_2, p-1)\nu(|\eta_1 - \eta_2|) \\ & + CH(\eta_1, \xi_1, \eta_2, \xi_2, p-1-s)|\xi_1 - \xi_2|^s, \end{aligned} \tag{3.5}$$

where $s > 0$, $p > 1$, $s \in (0, \min(p-1, 1))$ and ν is the modulus of continuity, a bounded, concave, and continuous function in \mathbb{R}_+ such that $\nu(0) = 0$, $\nu(t) = 1$ for $t \geq 1$ and $\nu(t) > 0$ for $t > 0$. Throughout the paper C and c (sometimes with indices) denote generic constants, q is defined by $1/p + 1/q = 1$, $y = x/\epsilon$, and $\|\cdot\|_{p,\Omega}$ denotes L^p -norm (either vector or scalar). In further analysis $K \in \mathbf{K}^h$ will be referred to simply by K . Inequalities (3.1)-(3.5) are the general conditions that guarantee the existence

of a solution and are used in homogenization of nonlinear operators [28]. Here p represents the rate of the polynomial growth of the fluxes with respect to gradient and, consequently, it controls the summability of the solution. We do not assume any differentiability with respect to η and ξ in the coefficients. Our objective is to present MsFEM and study its convergence for general nonlinear equations, where the fluxes can be discontinuous functions in space. These kinds of equations arise in many applications such as nonlinear heat conduction, nonlinear elasticity, flow in porous media, and etc. (see, e.g., [27, 32, 33, 24]). Our interest is in the applications to porous media flows related to flow in unsaturated media.

In [11] we have shown using G -convergence theory that

$$\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|u_h - u\|_{W_0^{1,p}(\Omega)} = 0, \tag{3.6}$$

(up to a subsequence) where u is a solution of (1.2) and u_h is a MsFEM solution given by (2.3). This result can be obtained without any assumption on the nature of the heterogeneities and can not be improved because there could be infinitely many scales, $\alpha(\epsilon)$, present such that $\alpha(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$.

For the periodic case (and general random homogeneous case) our goal is to show the convergence of MsFEM in the limit as $\epsilon/h \rightarrow 0$. To show the convergence for $\epsilon/h \rightarrow 0$, we consider $h = h(\epsilon)$, such that $h(\epsilon) \gg \epsilon$ and $h(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$. We would like to note that this limit as well as the proof of the periodic case is different from (3.6), where the double-limit is taken. In contrast to the proof of (3.6), the proof of the periodic case requires the correctors for the solutions of the local problems.

Next we will present the convergence results for MsFEM solutions. For general nonlinear elliptic equations under the assumptions (3.1)-(3.5) the strong convergence of MsFEM solutions can be shown. In the proof of this theorem we show the form of the truncation error (in a weak sense) in terms of the resonance errors between the mesh size and small scale ϵ . The resonance errors are derived explicitly. To obtain the convergence rate from the truncation error, one needs some lower bounds. Under the general conditions, such as (3.1)-(3.5), one can prove strong convergence of MsFEM solutions without an explicit convergence rate (cf.[33]). To convert the obtained convergence rates for the truncation errors into the convergence rate of MsFEM solutions, additional assumptions, such as monotonicity, are needed. This is discussed at the end of this section.

THEOREM 3.1. *Assume $a_\epsilon(x, \eta, \xi)$ and $a_{0,\epsilon}(x, \eta, \xi)$ are periodic functions with respect to x and let u be a solution of (1.2) and u_h is a MsFEM solution given by (2.3). Moreover, we assume that $D_x u_h$ is uniformly bounded in $L^{p+\alpha}(\Omega)$ for some $\alpha > 0$ ¹. Then*

$$\lim_{\epsilon \rightarrow 0} \|u_h - u\|_{W_0^{1,p}(\Omega)} = 0 \tag{3.7}$$

where $h = h(\epsilon) \gg \epsilon$ and $h \rightarrow 0$ as $\epsilon \rightarrow 0$ (up to a subsequence).

THEOREM 3.2. *Let u and u_h be the solutions of the homogenized problem (1.2) and MsFEM (2.3), respectively, with the coefficient $a_\epsilon(x, \eta, \xi) = a(x/\epsilon, \xi)$ and $a_{0,\epsilon} = 0$. Then*

$$\|u_h - u\|_{W_0^{1,p}(\Omega)}^p \leq c \left(\frac{\epsilon}{h}\right)^{\frac{s}{(p-1)(p-s)}} + c \left(\frac{\epsilon}{h}\right)^{\frac{p}{p-1}} + ch^{\frac{p}{p-1}}. \tag{3.8}$$

¹Please see Remark 3.1 at the end of the proof of Theorem 3.1 for more discussions and partial results regarding this assumption.

We will first prove Theorem 3.1. Then, using the estimates obtained in the proof of this theorem, we will show (3.8). The main idea of the proof of Theorem 3.1 is the following. First, the boundedness of the discrete solutions independent of ϵ and h will be shown. This allows us to extract a weakly converging sub-sequence. The next task is to prove that a limit is a solution of the homogenized equation. For this reason correctors for $v_{\epsilon,h}$ (see (2.2)) are used and their convergence is demonstrated. We would like to note that the known convergence results for the correctors assume a fixed (given) homogenized solution, while the correctors for $v_{\epsilon,h}$ are defined for only uniformly bounded sequence v_h , i.e., the homogenization limits of $v_{\epsilon,h}$ (with respect to ϵ) depend on h , and are only uniformly bounded. Because of this, more precise corrector results need to be obtained where the homogenized limit of the solution is tracked carefully in the analysis. Note that to prove (3.6) (see [13]), one does not need correctors and can use the fact of the convergence of fluxes, and, thus, the proof of the periodic case presented in this paper differs from the one in [13]. Some results of our paper (Lemmas 3.3, 3.4, and their proofs) do not require periodicity assumptions. For these results we will use the notations $a_\epsilon(x,\eta,\xi)$ and $a_{0,\epsilon}(x,\eta,\xi)$ to distinguish the two cases. The rest of the proofs require periodicity, and we will use $a(x/\epsilon,\eta,\xi)$ and $a_0(x/\epsilon,\eta,\xi)$ notations.

LEMMA 3.3. *There exists a constant $C > 0$ such that for any $v_h \in S^h$*

$$\langle A_{\epsilon,h} v_h, v_h \rangle \geq C \|D_x v_h\|_{p,\Omega}^p,$$

for sufficiently small h .

The proof of this lemma is provided in the Appendix 5.3. The following lemma will be used in the proof of Lemma 3.5.

LEMMA 3.4. *Let $v_\epsilon - v_0 \in W_0^{1,p}(K)$ and $w_\epsilon - w_0 \in W_0^{1,p}(K)$ satisfy the following problems, respectively:*

$$-\operatorname{div}(a_\epsilon(x,\eta,D_x v_\epsilon)) = 0 \text{ in } K \tag{3.9}$$

$$-\operatorname{div}(a_\epsilon(x,\eta,D_x w_\epsilon)) = 0 \text{ in } K \tag{3.10}$$

where η is constant in K . Then the following estimate holds:

$$\|D_x(v_\epsilon - w_\epsilon)\|_{p,K}^p \leq C H_0 \|D_x(v_0 - w_0)\|_{p,K}^{\frac{p}{p-s}}, \tag{3.11}$$

where

$$H_0 = \left(|K| + \|\eta\|_{p,K}^p + \|D_x v_0\|_{p,K}^p + \|D_x w_0\|_{p,K}^p \right)^{(p-s-1)/(p-s)},$$

where $s \in (0, \min(1, p-1))$, $p > 1$.

Proof of this lemma is presented in Appendix B.

Regarding η^{v_h} , where $\eta^{v_h} = \frac{1}{|K|} \int_K v_h dx$ in each K , we note that Jensen's inequality implies

$$\|\eta^{v_h}\|_{p,\Omega} \leq C \|v_h\|_{p,\Omega}. \tag{3.12}$$

In addition, the following estimates hold for η^{v_h} :

$$\|v_h - \eta^{v_h}\|_{p,K} \leq Ch \|D_x v_h\|_{p,K}. \tag{3.13}$$

At this stage we define a numerical corrector associated with $v_{\epsilon,h} = E^{MsFEM} v_h$, $v_h \in S^h$. First, let

$$P_{\eta,\xi}(y) = \xi + D_y N_{\eta,\xi}(y), \tag{3.14}$$

for $\eta \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$, where $N_{\eta,\xi} \in W_{per}^{1,p}(Y)$ is the periodic solution (with average zero) of

$$-\operatorname{div}(a(y,\eta,\xi + D_y N_{\eta,\xi}(y))) = 0 \text{ in } Y, \tag{3.15}$$

where Y is a unit period. The homogenized fluxes are defined as follows:

$$a^*(\eta,\xi) = \int_Y a(y,\eta,\xi + D_y N_{\eta,\xi}(y)) dy, \tag{3.16}$$

$$a_0^*(\eta,\xi) = \int_Y a_0(y,\eta,\xi + D_y N_{\eta,\xi}(y)) dy, \tag{3.17}$$

where a^* and a_0^* satisfy the conditions similar to (3.1) - (3.5). We refer to [28] for further details. Using (3.14), we denote our numerical corrector by \mathcal{P} which is defined as

$$\mathcal{P} = P_{\eta^{v_h}, D_x v_h} = D_x v_h + D_y N_{\eta^{v_h}, D_x v_h}(y). \tag{3.18}$$

Here η^{v_h} is a piece-wise constant function defined in each $K \in \mathbf{K}^h$ by $\eta^{v_h} = \frac{1}{|K|} \int_K v_h dx$. Consequently, \mathcal{P} is defined in Ω by using (3.18) in each $K \in \mathbf{K}^h$. For the linear problem $\mathcal{P} = D_x v_h + N(y) \cdot D_x v_h$. Our goal is to show the convergence of these correctors for a uniformly bounded family of v_h in $W^{1,p}(\Omega)$. We would like to note that the corrector results known in the literature are for a fixed homogenized solution.

LEMMA 3.5. *Let $v_{\epsilon,h}$ satisfy (2.2), where $a_\epsilon(x,\eta,\xi)$ is periodic function with respect to x , and assume that v_h is uniformly bounded in $W_0^{1,p}(\Omega)$. Then*

$$\|D_x v_{\epsilon,h} - \mathcal{P}\|_{p,\Omega} \leq C \left(\frac{\epsilon}{h}\right)^{\frac{1}{p(p-s)}} \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p\right)^{\frac{1}{p}}. \tag{3.19}$$

We note that here $s \in (0, \min(p-1, 1))$, $p > 1$. For the proof of this lemma, we need the following proposition.

PROPOSITION 3.6. *For every $\eta \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$ we have*

$$\|P_{\eta,\xi}\|_{p,Y_\epsilon}^p \leq c(1 + |\eta|^p + |\xi|^p) |Y_\epsilon|, \tag{3.20}$$

where Y_ϵ is a period of size ϵ . An easy consequence of this proposition is the following estimate for $N_{\eta,\xi}$ (see (3.15)).

COROLLARY 3.7. *For every $\eta \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$ we have*

$$\|D_y N_{\eta,\xi}\|_{p,Y_\epsilon}^p \leq c(1 + |\eta|^p + |\xi|^p) |Y_\epsilon|. \tag{3.21}$$

The proof of Proposition 3.6 is presented in Appendix C.

Proof. (Lemma 3.5) Recall that by definition $\mathcal{P} = D_x v_h + D_y N_{\eta^{v_h}, D_x v_h}(y) = D_x v_h + \epsilon D_x N_{\eta^{v_h}, D_x v_h}(x/\epsilon)$, where by using (3.15) $N_{\eta^{v_h}, D_x v_h}(y)$ is a zero-mean periodic function satisfying the following:

$$-\operatorname{div}(a(x/\epsilon, \eta^{v_h}, D_x v_h + \epsilon D_x N_{\eta^{v_h}, D_x v_h})) = 0 \text{ in } K. \tag{3.22}$$

We expand $v_{\epsilon,h}$ as

$$v_{\epsilon,h} = v_h(x) + \epsilon N_{\eta^{v_h}, D_x v_h}(x/\epsilon) + \theta(x, x/\epsilon). \tag{3.23}$$

We note that here $\theta(x, x/\epsilon)$ is similar to the correction terms that arise in linear problems because of the mismatch between linear boundary conditions and the oscillatory corrector, $N_{\eta^{v_h}, D_x v_h}(x/\epsilon) = N(x/\epsilon) \cdot D_x v_h$. Next we denote by $w_{\epsilon,h} = v_h(x) + \epsilon N_{\eta^{v_h}, D_x v_h}(x/\epsilon)$. Clearly $w_{\epsilon,h}$ satisfies (3.22). Taking all these into account, the claim in the lemma is the same as to proving

$$\|D_x \theta\|_{p,\Omega} = \|D_x(v_{\epsilon,h} - w_{\epsilon,h})\|_{p,\Omega} \leq C \left(\frac{\epsilon}{h}\right)^{\frac{1}{p(p-s)}} \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p\right)^{\frac{1}{p}}. \tag{3.24}$$

Here we may write $w_{\epsilon,h}$ as a solution of the following boundary value problem:

$$-\operatorname{div}(a(x/\epsilon, \eta^{v_h}, D_x w_{\epsilon,h})) = 0 \text{ in } K \quad \text{and} \quad w_{\epsilon,h} = v_h + \epsilon \tilde{N}_{\eta^{v_h}, D_x v_h} \text{ on } \partial K,$$

with $\tilde{N}_{\eta^{v_h}, D_x v_h} = \varphi N_{\eta^{v_h}, D_x v_h}$, where φ is a sufficiently smooth function whose value is 1 on a strip of width ϵ adjacent to ∂K and 0 elsewhere. We denote this strip by S_ϵ . This idea has been used in [22]. By Lemma 3.4 we have the following estimate:

$$\begin{aligned} \|D_x \theta\|_{p,K}^p &= \|D_x(v_{\epsilon,h} - w_{\epsilon,h})\|_{p,K}^p \\ &\leq C H_0 \|D_x(v_h - v_h - \epsilon \tilde{N}_{\eta^{v_h}, D_x v_h})\|_{p,K}^{\frac{p}{p-s}} \\ &\leq C H_0 \|\epsilon D_x \tilde{N}_{\eta^{v_h}, D_x v_h}\|_{p,K}^{\frac{p}{p-s}}, \end{aligned} \tag{3.25}$$

where

$$H_0 = \left(|K| + \|\eta^{v_h}\|_{p,K}^p + \|D_x v_h\|_{p,K}^p + \|D_x(v_h + \epsilon \tilde{N}_{\eta^{v_h}, D_x v_h})\|_{p,K}^p\right)^{\frac{p-s-1}{p-s}}. \tag{3.26}$$

We need to show that H_0 is bounded and $\|\epsilon D_x \tilde{N}_{\eta^{v_h}, D_x v_h}\|_{p,\Omega}^p$ uniformly vanishes as $\epsilon \rightarrow 0$. For this purpose, we use the following notations: let $J_\epsilon^K = \{i \in \mathbb{Z}^d : Y_\epsilon^i \cap K \neq \emptyset, K \setminus Y_\epsilon^i \neq \emptyset\}$ and $F_\epsilon^K = \cup_{i \in J_\epsilon^K} Y_\epsilon^i$. In other words, F_ϵ^K is the union of all periods Y_ϵ^i that cover the strip S_ϵ . Using these notations and because φ is zero everywhere in K , except in the strip S_ϵ , we may write the following:

$$\begin{aligned} \|\epsilon D_x \tilde{N}_{\eta^{v_h}, D_x v_h}\|_{p,K}^p &= \epsilon^p \int_K |D_x(\varphi N_{\eta^{v_h}, D_x v_h})|^p dx \\ &= \epsilon^p \int_{S_\epsilon} |D_x(\varphi N_{\eta^{v_h}, D_x v_h})|^p dx \\ &\leq \epsilon^p \int_{F_\epsilon^K} |D_x(\varphi N_{\eta^{v_h}, D_x v_h})|^p dx \\ &= \epsilon^p \sum_{i \in J_\epsilon^K} \int_{Y_\epsilon^i} |D_x(\varphi N_{\eta^{v_h}, D_x v_h})|^p dx \\ &\leq \epsilon^p \sum_{i \in J_\epsilon^K} \int_{Y_\epsilon^i} (|D_x N_{\eta^{v_h}, D_x v_h}|^p |\varphi|^p + |N_{\eta^{v_h}, D_x v_h}|^p |D_x \varphi|^p) dx, \end{aligned} \tag{3.27}$$

where we have used the product rule on the partial derivative in the last line of (3.27). Our aim now is to show that the sum of integrals in the last line of (3.27) is uniformly bounded. We note that (see Corollary 3.7)

$$\|D_y N_{\eta^{v_h}, D_x v_h}\|_{p, Y_\epsilon^i}^p \leq C(1 + |\eta^{v_h}|^p + |D_x v_h|^p) |Y_\epsilon^i|, \tag{3.28}$$

from which, using the Poincaré-Friedrich inequality we have

$$\|N_{\eta^{v_h}, D_x v_h}\|_{p, Y_\epsilon^i}^p \leq C(1 + |\eta^{v_h}|^p + |D_x v_h|^p) |Y_\epsilon^i|. \tag{3.29}$$

We note also that η^{v_h} and $D_x v_h$ are constant in K . Because φ is sufficiently smooth, and whose value is one on the strip S_ϵ and zero elsewhere, we know that $|D_x \varphi| \leq C/\epsilon$ (cf. [22]). Applying all these facts to (3.27) we have

$$\begin{aligned} \|\epsilon D_x \tilde{N}_{\eta^{v_h}, D_x v_h}\|_{p, K}^p &\leq C \epsilon^p (1 + |\eta^{v_h}|^p + |D_x v_h|^p) \sum_{i \in J_\epsilon^K} (1 + \epsilon^{-p}) |Y_\epsilon^i| \\ &= C (\epsilon^p + 1) (1 + |\eta^{v_h}|^p + |D_x v_h|^p) \sum_{i \in J_\epsilon^K} |Y_\epsilon^i| \\ &\leq C (1 + |\eta^{v_h}|^p + |D_x v_h|^p) \sum_{i \in J_\epsilon^K} |Y_\epsilon^i|. \end{aligned} \tag{3.30}$$

Moreover, because all Y_ϵ^i , $i \in J_\epsilon^K$, cover the strip S_ϵ , we know that $\sum_{i \in J_\epsilon^K} |Y_\epsilon^i| \leq C h^{d-1} \epsilon$. Hence, we have

$$\begin{aligned} \|\epsilon D_x \tilde{N}_{\eta^{v_h}, D_x v_h}\|_{p, K}^p &\leq C \frac{h^d}{h^d} (1 + |\eta^{v_h}|^p + |D_x v_h|^p) h^{d-1} \epsilon \\ &\leq C \frac{\epsilon}{h} \left(|K| + \|\eta^{v_h}\|_{p, K}^p + \|D_x v_h\|_{p, K}^p \right). \end{aligned} \tag{3.31}$$

Furthermore, using this estimate and noting that $\epsilon/h < 1$, we obtain from (3.26) that

$$H_0 \leq C \left(|K| + \|\eta^{v_h}\|_{p, K}^p + \|v_h\|_{p, K}^p + \|D_x v_h\|_{p, K}^p \right)^{\frac{p-s-1}{p-s}}. \tag{3.32}$$

Summarizing the results from (3.25) combined with (3.32) and (3.31), we get

$$\begin{aligned} \|D_x \theta\|_{p, K}^p &\leq C H_0 \|\epsilon D_x \tilde{N}_{\eta^{v_h}, D_x v_h}\|_{p, K}^{\frac{p}{p-s}} \\ &\leq C \left(\frac{\epsilon}{h} \right)^{\frac{1}{p-s}} \left(|K| + \|\eta^{v_h}\|_{p, K}^p + \|v_h\|_{p, K}^p + \|D_x v_h\|_{p, K}^p \right). \end{aligned} \tag{3.33}$$

Finally summing over all $K \in \mathbf{K}^h$ and applying (3.12) to $\sum_{K \in \mathbf{K}^h} \|\eta^{v_h}\|_{p, K}^p$, we obtain

$$\begin{aligned} \|D_x \theta\|_{p, \Omega}^p &= \sum_K \|D_x \theta\|_{p, K}^p \\ &\leq C \left(\frac{\epsilon}{h} \right)^{\frac{1}{p-s}} \sum_K \left(|K| + \|v_h\|_{p, K}^p + \|D_x v_h\|_{p, K}^p \right) \\ &= C \left(\frac{\epsilon}{h} \right)^{\frac{1}{p-s}} \left(|\Omega| + \|v_h\|_{p, \Omega}^p + \|D_x v_h\|_{p, \Omega}^p \right). \end{aligned} \tag{3.34}$$

The last inequality uniformly vanishes as ϵ approaching zero, thus we have completed the proof of Lemma 3.5. \square

The next lemma is crucial for the proof of Theorem 3.1 and it guarantees the convergence of MsFEM solutions to a solution of the homogenized equation. This lemma also provides us with the estimate for the truncation error (in a weak sense).

LEMMA 3.8. *Suppose $v_h, w_h \in S^h$ where $D_x v_h$ and $D_x w_h$ are uniformly bounded in $L^{p+\alpha}(\Omega)$ and $L^p(\Omega)$, respectively, for some $\alpha > 0$. Let A^* be the operator associated with the homogenized problem (1.2), such that*

$$\langle A^* v_h, w_h \rangle = \sum_{K \in \mathbf{K}^h} \int_K ((a^*(v_h, D_x v_h), D_x w_h) + a_0^*(v_h, D_x v_h) w_h) dx, \quad \forall v_h, w_h \in S^h. \tag{3.35}$$

Then we have

$$\lim_{\epsilon \rightarrow 0} \langle A_{\epsilon,h} v_h - A^* v_h, w_h \rangle = 0. \tag{3.36}$$

The proof of this lemma is presented in Appendix D. Now we are ready to prove Theorem 3.1.

Proof. (Theorem 3.1) Since $A_{\epsilon,h}$ is coercive, it follows that u_h is bounded, which implies that it has a subsequence (which we also denote by u_h) such that $u_h \rightharpoonup \tilde{u}$ in $W^{1,p}(\Omega)$ as $\epsilon \rightarrow 0$. Because the operator A^* is of type S_+ (see, e.g., [33], page 3, for the definition), then by its definition, the strong convergence would be true if we can show that $\limsup_{\epsilon \rightarrow 0} \langle A^* u_h, u_h - \tilde{u} \rangle \rightarrow 0$. Moreover, by adding and subtracting the term, we have the following equality:

$$\begin{aligned} \langle A^* u_h, u_h - \tilde{u} \rangle &= \langle A^* u_h - A_{\epsilon,h} u_h, u_h - \tilde{u} \rangle + \langle A_{\epsilon,h} u_h, u_h - \tilde{u} \rangle \\ &= \langle A^* u_h - A_{\epsilon,h} u_h, u_h \rangle - \langle A^* u_h - A_{\epsilon,h} u_h, \tilde{u} \rangle + \langle f, u_h - \tilde{u} \rangle. \end{aligned} \tag{3.37}$$

Lemma 3.8 implies that the first and second term vanish as $\epsilon \rightarrow 0$ provided $D_x u_h$ is uniformly bounded in $L^{p+\alpha}$ for $\alpha > 0$, while the last term vanishes as $\epsilon \rightarrow 0$ (up to a subsequence) by the weak convergence of u_h . One can assume additional mild regularity assumptions [26] for input data and obtain Meyers type estimates, $\|D_x u\|_{p+\alpha, \Omega} \leq C$, for the homogenized solutions. In this case it is reasonable to assume that the discrete solutions are uniformly bounded in $L^{p+\alpha}(\Omega)$. We have obtained results on Meyers type estimates for our approximate solutions in the case $p=2$ [12]. We are currently studying the generalizations of these results to arbitrary p . Finally, since A^* is also of type M (see, e.g., [32], page 38, for the definition), all these conditions imply that $A^* \tilde{u} = f$, which means that $\tilde{u} = u$. \square

REMARK 3.1. *We would like to point out that for the proof of Theorem 3.1 it is assumed that $D_x u_h$ is uniformly bounded in $L^{p+\alpha}(\Omega)$ for some $\alpha > 0$ (see, discussion after (3.37)). This has been shown for $p=2$ in [12]. To avoid this assumption, one can impose additional restrictions on $a^*(\eta, \xi)$ (see, [13], page 254–255). We would like to note that the assumption, $D_x u_h$ is uniformly bounded in $L^{p+\alpha}(\Omega)$, is not used for the estimation of the resonance errors, but only used in (D.12).*

Next we present some explicit estimates for the convergence rates of MsFEM. First, we note that from the proof of the Lemma 3.8 it follows that the truncation

error of MsFEM (in a weak sense) is given by

$$\begin{aligned}
& \langle A_{\epsilon,h}u_h - A^*u_h, w_h \rangle \\
&= \langle f - A^*u_h, w_h \rangle \leq c \left(\frac{\epsilon}{h} \right)^{\frac{s}{p(p-s)}} \left(|\Omega| + \|u_h\|_{p,\Omega}^p + \|D_x u_h\|_{p,\Omega}^p \right)^{\frac{1}{q}} \|D_x w_h\|_{p,\Omega} \\
&\quad + c \frac{\epsilon}{h} \left(|\Omega| + \|u_h\|_{p,\Omega}^p + \|D_x u_h\|_{p,\Omega}^p \right)^{\frac{1}{q}} \|D_x w_h\|_{p,\Omega} + e(h) \|D_x w_h\|_{p,\Omega} \\
&= c \left(\left(\frac{\epsilon}{h} \right)^{\frac{s}{p(p-s)}} + \frac{\epsilon}{h} \right) \left(|\Omega| + \|u_h\|_{p,\Omega}^p + \|D_x u_h\|_{p,\Omega}^p \right)^{\frac{1}{q}} \|D_x w_h\|_{p,\Omega} + e(h) \|D_x w_h\|_{p,\Omega},
\end{aligned} \tag{3.38}$$

where $e(h)$ is a generic sequence independent of small scale ϵ , such that $e(h) \rightarrow 0$ as $h \rightarrow 0$. In particular, the first, second, and third terms on the right side of (3.38) are the estimates of $\sum_{K \in \mathbf{K}^h} IK$, $\sum_{K \in \mathbf{K}^h} IIK$, and $\sum_{K \in \mathbf{K}^h} IIIK$, respectively (see (D.3) and (D.4)). We note that the first term on the right side of (3.38) is the leading order resonance error caused by the linear boundary conditions imposed on ∂K , the second term is due to mismatch between the mesh size and the small scale of the problem. These resonance errors are also present in the linear case [14]. If one uses the periodic solution of the auxiliary problem for constructing the solutions of the local problems, then the resonance error can be removed. To obtain explicit convergence rates, we first derive upper bounds for $\langle A^*u_h - A^*P_h u, u_h - P_h u \rangle$, where $P_h u$ denotes a finite element projection of u onto S^h , i.e.,

$$\langle A^*P_h u, v_h \rangle = \langle f, v_h \rangle, \quad \forall v_h \in S^h,$$

and $\langle A^*u_h, v_h \rangle$ is defined by (3.35). Then using estimate (3.38), we have

$$\begin{aligned}
& \langle A^*u_h - A^*P_h u, u_h - P_h u \rangle \\
&= \langle A^*u_h - A_{\epsilon,h}u_h, u_h - P_h u \rangle + \langle A_{\epsilon,h}u_h - A^*P_h u, u_h - P_h u \rangle \\
&= \langle A^*u_h - A_{\epsilon,h}u_h, u_h - P_h u \rangle + \langle f - A^*P_h u, u_h - P_h u \rangle \\
&= \langle A^*u_h - A_{\epsilon,h}u_h, u_h - P_h u \rangle \\
&\leq c \left(\left(\frac{\epsilon}{h} \right)^{\frac{s}{p(p-s)}} + \frac{\epsilon}{h} \right) \left(|\Omega| + \|u_h\|_{p,\Omega}^p + \|D_x u_h\|_{p,\Omega}^p \right)^{\frac{1}{q}} \\
&\quad \|D_x(u_h - P_h u)\|_{p,\Omega} + e(h) \|D_x(u_h - P_h u)\|_{p,\Omega}.
\end{aligned} \tag{3.39}$$

The estimate (3.39) does not allow us to obtain an explicit convergence rate without some lower bound for the left side of the expression. In the proof of Theorem 3.1, we only use the fact that A^* is the operator of type S_+ , which guarantees that the convergence of the left side of (3.39) to zero implies the convergence of the discrete solutions to a solution of the homogenized equation. Explicit convergence rates can be obtained by assuming some kind of an inverse stability condition, $\|A^*u - A^*v\| \geq c\|u - v\|$. In particular, we may assume that A^* is a monotone operator, i.e.,

$$\langle A^*u - A^*v, u - v \rangle \geq c \|D_x(u - v)\|_{p,\Omega}^p. \tag{3.40}$$

A simple way to achieve monotonicity is to assume $a_\epsilon(x, \eta, \xi) = a_\epsilon(x, \xi)$ and $a_{0,\epsilon}(x, \eta, \xi) = 0$, though one can impose additional conditions on $a_\epsilon(x, \eta, \xi)$ and

$a_{0,\epsilon}(x, \eta, \xi)$, such that monotonicity condition (3.40) is satisfied. For our further calculations, we only assume (3.40). Then from (3.39) and (3.40), and using Young inequality, we have

$$\|D_x(u_h - P_h u)\|_{p,\Omega}^p \leq c \left(\frac{\epsilon}{h}\right)^{\frac{s}{(p-1)(p-s)}} + c \left(\frac{\epsilon}{h}\right)^{\frac{p}{p-1}} + e(h).$$

Next taking into account the convergence of standard finite element solutions of the homogenized equation we write

$$\|D_x P_h u - D_x u\|_{p,\Omega} \leq e_1(h),$$

where $e_1(h) \rightarrow 0$ (as $h \rightarrow 0$) is independent of ϵ . Consequently, using triangle inequality we have

$$\|D_x(u_h - u)\|_{p,\Omega}^p \leq c \left(\frac{\epsilon}{h}\right)^{\frac{s}{(p-1)(p-s)}} + c \left(\frac{\epsilon}{h}\right)^{\frac{p}{p-1}} + e(h) + e_1(h).$$

Proof. (Theorem 3.2). For monotone operators, $a_\epsilon(x, \eta, \xi) = a_\epsilon(x, \xi)$ and $a_{0,\epsilon}(x, \eta, \xi) = 0$, $\eta \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$, the estimates for $e(h)$ and $e_1(h)$ can be easily derived. In particular, because of the absence of η in a_ϵ , $e(h) = 0$ (see (D.3) and (D.4)), while $e_1(h) \leq Ch^{\frac{1}{p-1}}$ (see for example [7]). Combining these estimates we have

$$\|D_x(u_h - u)\|_{p,\Omega}^p \leq c \left(\frac{\epsilon}{h}\right)^{\frac{s}{(p-1)(p-s)}} + c \left(\frac{\epsilon}{h}\right)^{\frac{p}{p-1}} + ch^{\frac{p}{p-1}}. \tag{3.41}$$

From here one obtains (3.8). \square

REMARK 3.2. One can impose various conditions on the operators to obtain different kinds of convergence rates. For example, under the additional assumptions (cf. [27])

$$\left| \frac{\partial a^*(\eta, \xi)}{\partial \eta} \right| + \left| \frac{\partial a^*(\eta, \xi)}{\partial \xi} \right| \leq C, \quad \frac{\partial a_i^*(\eta, \xi)}{\partial \xi_j} \zeta_i \zeta_j \geq C|\zeta|^2, \tag{3.42}$$

where $\zeta \in \mathbb{R}^d$ is an arbitrary vector, and $p=2$, following the analysis presented in [27], pages 51–52, the convergence rate in terms of L^p -norm of $u_h - P_h u$ can be obtained,

$$\|D_x(u_h - P_h u)\|_{p,\Omega}^p \leq c \left(\frac{\epsilon}{h}\right)^{\frac{s}{(p-1)(p-s)}} + c \left(\frac{\epsilon}{h}\right)^{\frac{p}{p-1}} + e(h) + C\|u_h - P_h u\|_{p,\Omega}^p,$$

where $s \in (0, 1)$, $p=2$.

REMARK 3.3. For the linear operators ($p=2$, $s=1$), we recover the convergence rate obtained in [20], $Ch + C_1\sqrt{\epsilon/h}$.

REMARK 3.4. We have shown that MsFEM for nonlinear problems has the same error structure as for the linear problems. In particular, our studies revealed two kinds of resonance errors for nonlinear problems with the same nature as those that arise in linear problems [14].

3.1. Approximation of the oscillations. In this section we present a theorem demonstrating the approximation of oscillatory solutions u_ϵ of (1.1).

THEOREM 3.9. Let u_ϵ be the solution of boundary value problem (1.1) and $u_h \in S^h$ and $u_{\epsilon,h} \in V_\epsilon^h$ with $u_{\epsilon,h} = E^{MsFEM} u_h$ be MsFEM solution (homogenized and fluctuating components, respectively) (2.3). Then $\lim_{\epsilon \rightarrow 0} \|D_x u_{\epsilon,h} - D_x u_\epsilon\|_{p,\Omega} = 0$.

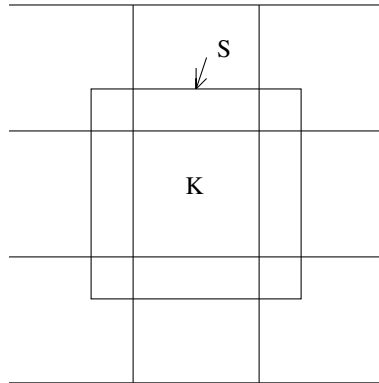


FIG. 3.1. Illustration of oversampling domain. K is a target coarse block, S is an oversampled domain

Below, we outline the proof (the details can be found in [17]). For the proof of this theorem we first introduce correctors for u_ϵ . Define an operator approximating the identity map in $L^p(\Omega)$ by

$$M_\epsilon \varphi(x) = \sum_{i \in I_\epsilon} \Psi_{Y_\epsilon^i}(x) \frac{1}{|Y_\epsilon^i|} \int_{Y_\epsilon^i} \varphi(y) dy, \tag{3.43}$$

where Y_ϵ^i is a period of size ϵ for $i \in \mathbb{Z}^d$, $I_\epsilon = \{i \in \mathbb{Z}^d : Y_\epsilon^i \subset \Omega\}$ and $\Psi_{Y_\epsilon^i}$ is an indicator function for Y_ϵ^i . Next we denote $P = P_{M_\epsilon u, M_\epsilon D_x u}(x, x/\epsilon) = M_\epsilon D_x u(x) + D_y N_{M_\epsilon u, M_\epsilon D_x u}(x/\epsilon)$, where u is the solution of the homogenized problem. The function P is a corrector associated with the original boundary value problem (1.1). Now, by triangle inequality, we have

$$\|D_x u_{\epsilon, h} - D_x u_\epsilon\|_{p, \Omega} \leq \|D_x u_{\epsilon, h} - \mathcal{P}\|_{p, \Omega} + \|\mathcal{P} - P\|_{p, \Omega} + \|P - D_x u_\epsilon\|_{p, \Omega}, \tag{3.44}$$

where $\mathcal{P} = D_x u_h + D_y N_{\eta^h, D_x u_h}(y)$ as defined before for $v_{\epsilon, h}$. Lemma 3.5 gives the convergence of the first term. The convergence of the third term can be obtained using the techniques developed in [8] and the details are in [17]. The convergence of the second term is due to the fact that $D_x u_h \rightarrow D_x u$ in $L^p(\Omega)^d$, and the details can be found in [17].

3.2. An Oversampling Technique. The approximation property of the corrector $\mathcal{P}(x, x/\epsilon)$ (cf. Lemma 3.5) reveals the existence of a resonance error proportional to ϵ/h , which is resulted from the mismatch due to the imposed linear boundary conditions for the local problem in the multiscale map E^{MsFEM} . The correction that arises in the expansion of the local solutions (3.23) and accounts for the mismatch is given by $\theta(x, x/\epsilon)$. As in the linear problems ([19]), $\theta(x, x/\epsilon)$ is the cause of the resonance errors. In [19] the authors suggested an oversampling technique for reducing the effects of $\theta(x, x/\epsilon)$. The main idea of the oversampling technique is to use the solutions of the local problems (cf. (2.2)) in larger domains. Here we extend this technique to nonlinear equations. In particular, the multiscale map E^{MsFEM} is constructed using the solutions of the local problems on the element larger than h . For periodic problems, the size of the larger domain can be chosen to be $h + \epsilon$, though for

more general problems without scale separation the size of the larger domain can be chosen to be $h + \beta h$, where β is a constant. In our simulations for general anisotropic heterogeneities, we choose $\beta = 1$. Furthermore, only the information from the target element is used in the multiscale formulation of the problem (see Figure 3.1).

In general, given $v_h \in S^h$, where v_h is defined in K , we want to find $v_{\epsilon,h}$ that satisfies

$$-\operatorname{div}(a_\epsilon(x, \eta^{v_h}, D_x v_{\epsilon,h})) = 0 \quad \text{in } S \tag{3.45}$$

such that $v_{\epsilon,h}(z_i) = v_h(z_i)$, where z_i are the nodal points of the target coarse element K . Thus, in general, we need to find a solution of the local equation with given values at the nodal (interior) points. This problem can be solved for linear problems using linear combinations of the local solutions in larger domains S . Here we present an oversampling technique for special cases in which the gradient in the coefficient is linear, i.e., $a_\epsilon(x, \eta, \xi) = a(x/\epsilon, \eta)\xi$, given $v_h \in S^h$, we define

$$v_{\epsilon,h} = \sum_{i=1}^3 c_i \phi_\epsilon^i, \tag{3.46}$$

where ϕ_ϵ^i satisfies

$$\begin{aligned} -\operatorname{div}(a(x/\epsilon, \eta^{v_h}) D_x \phi_\epsilon^i) &= 0 \quad \text{in } S \\ \phi_\epsilon^i &= \phi_0^i \quad \text{on } \partial S. \end{aligned} \tag{3.47}$$

The constants c_i , $i = 1, 2, 3$ are determined by imposing the conditions

$$v_{\epsilon,h}(z_j) = v_h(z_j) \quad j = 1, 2, 3. \tag{3.48}$$

We note that the piecewise constants in η^{v_h} are taken as the average over the element K . We would like to note that the convergence analysis of MsFEM with an oversampling technique requires some modifications of the proof presented in this paper for MsFEM without oversampling. In particular, the improved corrector results (see Lemma 3.5) are necessary to show the advantages of MsFEM with oversampling. This is a subject of our future research.

4. Numerical results and applications

4.1. Oversampling vs. non-oversampling. In this section we present several ingredients pertaining to the implementation of the multiscale finite element method. We will present numerical results for both MsFEM and the multiscale finite volume element method (MsFVEM). We use an Inexact-Newton algorithm as an iterative technique to tackle the nonlinearity. For the numerical examples below, we use $a_\epsilon(x, u_\epsilon, D_x u_\epsilon) = a_\epsilon(x, u_\epsilon) D_x u_\epsilon$. Let $\{\phi_0^i\}_{i=1}^{N_{dof}}$ be the standard piecewise linear basis functions of S^h . Then the MsFEM solution may be written as

$$u_h = \sum_{i=1}^{N_{dof}} \alpha_i \phi_0^i \tag{4.1}$$

for some $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{N_{dof}})^T$, where α_i depends on ϵ . Hence, we need to find α such that

$$F(\alpha) = 0, \tag{4.2}$$

TABLE 4.1. *Relative MsFEM Errors without Oversampling*

N	L^2 -norm		H^1 -norm		L^∞ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.029		0.115		0.03	
64	0.053	-0.85	0.156	-0.44	0.0534	-0.94
128	0.10	-0.94	0.234	-0.59	0.10	-0.94

where $F : \mathbb{R}^{N_{dof}} \rightarrow \mathbb{R}^{N_{dof}}$ is a nonlinear operator such that

$$F_i(\alpha) = \sum_{K \in \mathbf{K}^h} \int_K (a_\epsilon(x, \eta^{u_h}) D_x u_{\epsilon,h}, D_x \phi_0^i) dx - \int_\Omega f \phi_0^i dx. \tag{4.3}$$

We note that in (4.3) α is implicitly buried in η^{u_h} and $u_{\epsilon,h}$. An inexact-Newton algorithm is a variation of Newton’s iteration for a nonlinear system of equations, where the Jacobian system is only approximately solved. To be specific, given an initial iterate α^0 , for $k = 0, 1, 2, \dots$ until convergence do the following:

- Solve $F'(\alpha^k) \delta^k = -F(\alpha^k)$ by some iterative technique until $\|F(\alpha^k) + F'(\alpha^k) \delta^k\| \leq \beta_k \|F(\alpha^k)\|$.
- Update $\alpha^{k+1} = \alpha^k + \delta^k$.

In this algorithm $F'(\alpha^k)$ is the Jacobian matrix evaluated at iteration k . We note that when $\beta_k = 0$ then we have recovered the classical Newton iteration. Here we have used

$$\beta_k = 0.001 \left(\frac{\|F(\alpha^k)\|}{\|F(\alpha^{k-1})\|} \right)^2, \tag{4.4}$$

with $\beta_0 = 0.001$. Choosing β_k this way, we avoid over-solving the Jacobian system when α^k is still considerably far from the exact solution.

Next we present the entries of the Jacobian matrix. For this purpose, we use the following notations. Let $\mathbf{K}_i^h = \{K \in \mathbf{K}^h : z_i \text{ is a vertex of } K\}$, $I^i = \{j : z_j \text{ is a vertex of } K \in \mathbf{K}_i^h\}$, and $\mathbf{K}_{ij}^h = \{K \in \mathbf{K}_i^h : K \text{ shares } \overline{z_i z_j}\}$. We note that we may write $F_i(\alpha)$ as follows:

$$F_i(\alpha) = \sum_{K \in \mathbf{K}_i^h} \left(\int_K (a_\epsilon(x, \eta^{u_h}) D_x u_{\epsilon,h}, D_x \phi_0^i) dx - \int_K f \phi_0^i dx \right), \tag{4.5}$$

with

$$-\text{div}(a_\epsilon(x, \eta^{u_h}) D_x u_{\epsilon,h}) = 0 \text{ in } K \quad \text{and} \quad u_{\epsilon,h} = \sum_{z_m \in Z_K} \alpha_m \phi_0^m \text{ on } \partial K, \tag{4.6}$$

where Z_K is all the vertices of element K . It is apparent that $F_i(\alpha)$ is not fully dependent on all $\alpha_1, \alpha_2, \dots, \alpha_d$. Consequently, $\frac{\partial F_i(\alpha)}{\partial \alpha_j} = 0$ for $j \notin I^i$. To this end, we denote $\psi_\epsilon^j = \frac{\partial u_{\epsilon,h}}{\partial \alpha_j}$. By applying the chain rule of differentiation to (4.6) we have the following local problem for ψ_ϵ^j :

$$-\text{div}(a_\epsilon(x, \eta^{u_h}) D_x \psi_\epsilon^j) = \frac{1}{3} \text{div} \left(\frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} D_x u_{\epsilon,h} \right) \text{ in } K \quad \text{and} \quad \psi_\epsilon^j = \phi_\epsilon^j \text{ on } \partial K. \tag{4.7}$$

TABLE 4.2. *Relative MsFVEM Errors without Oversampling*

.9 N	L^2 -norm		H^1 -norm		L^∞ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.03		0.13		0.04	
64	0.05	-0.65	0.19	-0.60	0.05	-0.24
128	0.058	-0.19	0.25	-0.35	0.057	-0.19

TABLE 4.3. *Relative MsFEM Errors with Oversampling*

N	L^2 -norm		H^1 -norm		L^∞ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.0016		0.036		0.0029	
64	0.0012	0.38	0.019	0.93	0.0016	0.92
128	0.0024	-0.96	0.0087	1.14	0.0026	-0.71

The fraction $1/3$ comes from taking the derivative in the chain rule of differentiation. In the formulation of the local problem, we have replaced the nonlinearity in the coefficient by η^{u_h} , where for each triangle K $\eta^{u_h} = 1/3 \sum_{i=1}^3 \alpha_i^K$, which gives $\partial \eta^{u_h} / \partial \alpha_i = 1/3$. Moreover, for a rectangular element the fraction $1/3$ should be replaced by $1/4$.

Thus, provided that $v_{\epsilon,h}$ has been computed, then we may compute ψ_ϵ^j using (4.7). Using the above descriptions we have the expressions for the entries of the Jacobian matrix:

$$\frac{\partial F_i}{\partial \alpha_i} = \sum_{K \in \mathbf{K}_i^h} \left(\frac{1}{3} \int_K \left(\frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} D_x u_{\epsilon,h}, D_x \phi_0^i \right) dx + \int_K (a_\epsilon(x, \eta^{u_h}) D_x \psi_i, D_x \phi_0^i) dx, \right) \tag{4.8}$$

$$\frac{\partial F_i}{\partial \alpha_j} = \sum_{K \in \mathbf{K}_{ij}^h} \left(\frac{1}{3} \int_K \left(\frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} D_x u_{\epsilon,h}, D_x \phi_\epsilon^i \right) dx + \int_K (a_\epsilon(x, \eta^{u_h}) D_x \psi_\epsilon^j, D_x \phi_0^i) dx, \right) \tag{4.9}$$

for $j \neq i, j \in I^i$.

The implementation of the oversampling technique is similar to the procedure presented above, except the local problems in larger domains are used. From (3.46), (3.47), and (3.48) we obtain $v_{\epsilon,h}$ that satisfies the homogeneous local problem. As in the non-oversampling case, we denote $\psi_\epsilon^j = \frac{\partial v_{\epsilon,h}}{\partial \alpha_j}$, such that after applying the chain rule of differentiation to the local problem we have:

$$-\operatorname{div}(a_\epsilon(x, \eta^{u_h}) D_x \psi_\epsilon^j) = \frac{1}{3} \operatorname{div} \left(\frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} D_x v_{\epsilon,h} \right) \text{ in } S \quad \text{and} \quad \psi_\epsilon^j = \phi_0^j \text{ on } \partial S, \tag{4.10}$$

where η^{u_h} is computed over the corresponding element K and ϕ_0^j is understood as the nodal basis functions on oversampled domain S . Then all the rest of the inexact-Newton algorithms are the same as in the non-oversampling case. Specifically, we also use (4.8) and (4.9) to construct the Jacobian matrix of the system. We note that we will only use ψ_ϵ^j from (4.10) pertaining to the element K .

TABLE 4.4. *Relative MsFVEM Errors with Oversampling*

N	L^2 -norm		H^1 -norm		L^∞ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.002		0.038		0.005	
64	0.003	-0.43	0.021	0.87	0.003	0.72
128	0.001	1.10	0.009	1.09	0.001	1.08

TABLE 4.5. *Relative MsFEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$*

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0505		0.0025		0.025	
64	0.0002	1.58	0.029	0.8	0.001	1.32	0.017	0.57
128	0.0001	1	0.016	0.85	0.0005	1	0.011	0.62

From the derivation (both for oversampling and non-oversampling) it is obvious that the Jacobian matrix is not symmetric but sparse. Computation of this Jacobian matrix is similar to computing the stiffness matrix resulting from a standard finite element, where each entry is formed by accumulation of element by element contribution. Once we have the matrix stored in memory, then its action to a vector is straightforward. Because it is a sparse matrix, devoting some amount of memory for entries storage is inexpensive. The resulting linear system is solved using the preconditioned bi-conjugate gradient stabilized method.

We want to solve the following problem:

$$\begin{aligned} -\operatorname{div}(a(x/\epsilon, u_\epsilon) D_x u_\epsilon) &= -1 && \text{in } \Omega \subset \mathbb{R}^2, \\ u_\epsilon &= 0 && \text{on } \partial\Omega, \end{aligned} \quad (4.11)$$

where $\Omega = [0, 1] \times [0, 1]$, $a(x/\epsilon, u_\epsilon) = k(x/\epsilon)/(1 + u_\epsilon)^{l(x/\epsilon)}$, with

$$k(x/\epsilon) = \frac{2 + 1.8 \sin(2\pi x_1/\epsilon)}{2 + 1.8 \cos(2\pi x_2/\epsilon)} + \frac{2 + \sin(2\pi x_2/\epsilon)}{2 + 1.8 \cos(2\pi x_1/\epsilon)} \quad (4.12)$$

and $l(x/\epsilon)$ is generated from $k(x/\epsilon)$ such that the average of $l(x/\epsilon)$ over Ω is 2. Here we use $\epsilon = 0.01$. Because the exact solution for this problem is not available, we use a well resolved numerical solution using the standard finite element method as a reference solution. The resulting nonlinear system is solved using an inexact-Newton algorithm. The reference solution is solved on 512×512 mesh. Tables 4.1 and 4.3 present the relative errors of the solution with and without oversampling, respectively. In tables 4.2 and 4.4, the relative errors for the multiscale finite volume element method are presented. The relative errors are computed as the corresponding error divided by the norm of the solution. In each table, the second, third, and fourth columns list the relative error in L^2 , H^1 , and L^∞ norm, respectively. As we can see from these two tables, the oversampling significantly improves the accuracy of the multiscale method.

For our next example, we consider the problem with non-periodic coefficients, where $a_\epsilon(x, \eta) = k_\epsilon(x)/(1 + \eta)^{\alpha_\epsilon(x)}$. $k_\epsilon(x) = \exp(\beta_\epsilon(x))$ is chosen such that $\beta_\epsilon(x)$ is a

TABLE 4.6. *Relative MsFVEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$*

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0515		0.0025		0.027	
64	0.0002	1.58	0.029	0.81	0.0013	0.94	0.018	0.58
128	0.0001	1	0.016	0.85	0.0005	1.38	0.012	0.58

realization of a random field with the spherical variogram [9] and with the correlation lengths $l_x = 0.2$, $l_y = 0.02$ and with the variance $\sigma = 1$. $\alpha_\epsilon(x)$ is chosen such that $\alpha_\epsilon(x) = k_\epsilon(x) + \text{const}$ with the spatial average of 2. As for the boundary conditions we use “left-to-right flow” in $\Omega = [0, 5] \times [0, 1]$ domain, $u_\epsilon = 1$ at the inlet ($x_1 = 0$), $u_\epsilon = 0$ at the outlet ($x_1 = 5$), and no flow boundary conditions on the lateral sides $x_2 = 0$ and $x_2 = 1$. In Table 4.5 we present the relative error for the multiscale method with oversampling. Similarly, in Table 4.6 we present the relative error for the multiscale finite volume method with oversampling. Clearly, the oversampling method captures the effects induced by the large correlation features. Both H^1 and horizontal flux errors are under five percent. Similar results have been observed for various kinds of non-periodic heterogeneities. In the next set of numerical examples, we test MsFEM for problems with fluxes $a_\epsilon(x, \eta)$ that are discontinuous in space. The discontinuity in the fluxes is introduced by multiplying the underlying permeability function, $k_\epsilon(x)$, by a constant in certain regions, while leaving it unchanged in the rest of the domain. As an underlying permeability field, $k_\epsilon(x)$, we choose the random field used for the results in Table 4.5. In the first set of examples, the discontinuities are introduced along the boundaries of the coarse elements. In particular, $k_\epsilon(x)$ on the left half of the domain is multiplied by a constant J , where $J = \exp(1)$, or $\exp(2)$, or $\exp(4)$. The results in Tables 4.7-4.9 show that MsFEM converges and the error falls below five percent for relatively large coarsening. For the second set of examples (Tables 4.10-4.12), the discontinuities are not aligned with the boundaries of the coarse elements. In particular, the discontinuity boundary is given by $y = x\sqrt{2} + 0.5$, i.e., the discontinuity line intersects the coarse grid blocks. Similar to the aligned case, various jump magnitudes are considered. These results demonstrate the robustness of our approach for anisotropic fields where h and ϵ are nearly the same, and the fluxes that are discontinuous spatial functions.

As for CPU comparisons, we have observed more than 92 percent CPU savings when using MsFEM without oversampling. With the oversampling approach, the CPU savings depend on the size of the oversampled domain. For example, if the oversampled domain size is two times larger than the target coarse block (half coarse block extension on each side) we have observed 70 percent CPU savings for 64×64 and 80 percent CPU savings for 128×128 coarse grid. In general, the computational cost will decrease if the oversampled domain size is close to the target coarse block size, and this cost will be close to the cost of MsFEM without oversampling. Conversely, the error decreases if the size of the oversampled domains increases. In the numerical examples studied in our paper, we have observed the same errors for the oversampling methods using either one coarse block extension or half coarse block extensions. The latter indicates that the leading resonance error is eliminated by using a smaller oversampled domain. Oversampled domains with one coarse block extension are pre-

TABLE 4.7. *Relative MsFEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$, aligned discontinuity, $jump = \exp(1)$*

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0641		0.0020		0.039	
64	0.0002	1.58	0.0382	0.75	0.0010	1.00	0.027	0.53
128	0.0001	1.00	0.0210	0.86	0.0005	1.00	0.018	0.59

TABLE 4.8. *Relative MsFEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$, aligned discontinuity, $jump = \exp(2)$*

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0008		0.0817		0.0040		0.061	
64	0.0004	1.00	0.0493	0.73	0.0023	0.80	0.041	0.57
128	0.0002	1.00	0.0256	0.95	0.0011	1.06	0.025	0.71

viously used in simulations of flow through heterogeneous porous media [36]. As it is indicated in [19], one can use large oversampled domains for simultaneous computations of the several local solutions. Moreover, parallel computations will improve the speed of the method because MsFEM is well suited for parallel computation [19]. For the problems where $a_\epsilon(x, \eta, \xi) = a_\epsilon(x)b(\eta)\xi$ (see section 2.2 and the next section for applications) our multiscale computations are very fast because the base functions are built in the beginning of the computations. In this case, we have observed more than 95 percent CPU savings.

4.2. Applications of MsFEM to Richards' equation. In this section we consider the applications of MsFEM to Richards' equation, which describes the infiltration of water flow into a porous media whose pore space is filled with air and some water. The equation describing Richards' equation under some assumptions is given by

$$D_t \theta(u) - \operatorname{div}(K(x, u) D_x(u + x_3)) = 0 \quad \text{in } \Omega, \quad (4.13)$$

where $\theta(u)$ is volumetric water content and u is the pressure. The followings are assumed ([30]) for (4.13): (1) the porous medium and water are incompressible; (2) the temporal variation of the water saturation is significantly larger than the temporal variation of the water pressure; (3) air phase is infinitely mobile so that the air pressure remains constant, in this case it is atmospheric pressure which equals zero; (4) neglect the source/sink terms.

Constitutive relations between θ and u and between K and u are developed appropriately, which consequently gives nonlinearity behavior in (4.13). The relation between the water content and pressure is referred to as the moisture retention function. The equation written in (4.13) is called the *coupled-form* of Richards' Equation. In other literature this equation is also called the mixed form of Richards' Equation, due to the fact that there are two variables involved in it, namely, the water content θ and the pressure head u . Taking advantage of the differentiability of the soil retention

TABLE 4.9. Relative MsFEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$, aligned discontinuity, $jump = \exp(4)$

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0011		0.1010		0.0068		0.195	
64	0.0006	0.87	0.0638	0.66	0.0045	0.59	0.109	0.84
128	0.0003	1.00	0.0349	0.87	0.0024	0.91	0.063	0.79

TABLE 4.10. Relative MsFEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$, nonaligned discontinuity, $jump = \exp(1)$

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0623		0.0023		0.035	
64	0.0002	1.58	0.0366	0.77	0.0014	0.72	0.024	0.54
128	0.0001	1.00	0.0203	0.85	0.0006	1.22	0.016	0.59

function, one may rewrite (4.13) as follows:

$$C(u)D_t u - \operatorname{div}(K(x, u)D_x(u + x_3)) = 0 \quad \text{in } \Omega, \tag{4.14}$$

where $C(u) = d\theta/du$ is the specific moisture capacity. This version is referred to as the *head-form* (*h-form*) of Richards' Equation. Another formulation of the Richards' Equation is based on the water content θ ,

$$D_t \theta - \operatorname{div}(D(x, \theta)D_x \theta) - \frac{\partial K}{\partial x_3} = 0 \quad \text{in } \Omega, \tag{4.15}$$

where $D(\theta) = K(\theta)/(d\theta/du)$ defines the diffusivity. This form is called the *θ -form* of Richards' Equation.

The sources of nonlinearity of Richards' Equation comes from the moisture retention and relative hydraulic conductivity functions, $\theta(u)$ and $K(x, u)$, respectively. Reliable approximation of these relations are in general tedious to develop and thus also challenging. Field measurements or laboratory experiments to gather the parameters are relatively expensive, and furthermore, even if one can come up with such relations from these works, they will be somehow limited to the particular cases under consideration.

Perhaps the most widely used empirical constitutive relations for the moisture content and hydraulic conductivity is due to the work of van Genuchten [34]. He proposed a method of determining the functional relation of relative hydraulic conductivity to the pressure head by using the field observation knowledge of the moisture retention. In turn, the procedure would require curve-fitting to the proposed moisture retention function with the experimental/observational data to establish certain parameters inherent to the resulting hydraulic conductivity model. There are several widely known formulations of the constitutive relations: **Haverkamp model** - $\theta(u) = \frac{\alpha(\theta_s - \theta_r)}{\alpha + |u|^\beta} + \theta_r$, $K(x, u) = K_s(x) \frac{A}{A + |u|^\gamma}$; **van Genuchten model** [34] - $\theta(u) = \frac{\alpha(\theta_s - \theta_r)}{[1 + (\alpha|u|)^n]^m} + \theta_r$, $K(x, u) = K_s(x) \frac{\{1 - (\alpha|u|)^{n-1} [1 + (\alpha|u|)^n]^{-m}\}^2}{[1 + (\alpha|u|)^n]^{m/2}}$; **Exponential model** [35] - $\theta(u) = \theta_s e^{\beta u}$, $K(x, u) = K_s(x) e^{\alpha u}$.

TABLE 4.11. *Relative MsFEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$, nonaligned discontinuity, $\text{jump} = \exp(2)$*

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0010		0.0785		0.0088		0.052	
64	0.0003	1.74	0.0440	0.84	0.0052	0.76	0.031	0.75
128	0.0001	1.59	0.0239	0.88	0.0022	1.24	0.017	0.87

TABLE 4.12. *Relative MsFEM Errors for random heterogeneities, spherical variogram, $l_x = 0.20$, $l_z = 0.02$, $\sigma = 1.0$, nonaligned discontinuity, $\text{jump} = \exp(4)$*

N	L^2 -norm		H^1 -norm		L^∞ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0067		0.1775		0.1000		0.164	
64	0.0016	2.07	0.0758	1.23	0.0288	1.80	0.077	1.09
128	0.0009	0.83	0.0687	0.14	0.0423	-0.55	0.039	0.98

The variable K_s in the above models is also known as the saturated hydraulic conductivity. It has been observed that the hydraulic conductivity has a broad range of values, which together with the functional forms presented above, confirm the nonlinear behavior of the process. Furthermore, the water content and hydraulic conductivity approach zero as the pressure head goes to very large negative values. In other words, the Richards' Equation has a tendency to degenerate in a very dry condition, i.e., conditions with the large negative pressure. Because we are interested in mass conservative schemes, finite volume formulation (2.5) of the global problem instead of finite element formulation will be used. For (4.13), it is to find $u^h \in S^h$ such that

$$\int_{V_z} (\theta(\eta^{u^h}) - \theta^{n-1}) dx - \Delta t \int_{\partial V_z} K(x, \eta^{u^h}) D_x u_{\epsilon, h} \cdot n dl = 0 \quad \forall z \in Z_h^0, \quad (4.16)$$

where θ^{n-1} is the value of $\theta(\eta^{u^h})$ evaluated at time step $n-1$, and $u_{\epsilon, h} \in V_\epsilon^h$ is a function that satisfies the boundary value problem:

$$\begin{aligned} -\text{div}(K(x, \eta^{u^h}) D_x u_{\epsilon, h}) &= 0 && \text{in } K \in S^h, \\ u_{\epsilon, h} &= u_h && \text{on } \partial K. \end{aligned} \quad (4.17)$$

Here V_z is the control volume surrounding the vertex $z \in Z_h^0$ and Z_h^0 is the collection of all vertices that do not belong to the Dirichlet boundary.

MsFEM (or MsFVEM) offers great advantage when the nonlinearity and heterogeneity of $K(x, p)$ is separable, i.e.,

$$K(x, u) = k_s(x) k_r(u). \quad (4.18)$$

In this case, as we discussed earlier, the local problems become linear and the corresponding V_ϵ^h is a linear space, i.e., we may construct a set of basis functions $\{\psi_z\}_{z \in Z_h^0}$ such that they satisfy

$$\begin{aligned} -\text{div}(k_s(x) D_x \psi_z) &= 0 && \text{in } K \in S^h, \\ \psi_z &= \phi_z && \text{on } \partial K, \end{aligned} \quad (4.19)$$

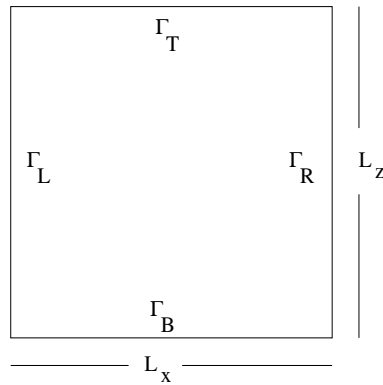


FIG. 4.1. Rectangular porous medium

where ϕ_z is a piecewise linear function. We note that if u_h has a discontinuity or a sharp front region, then the multiscale basis functions need to be updated only in that region. The latter is similar to the use of MsFEM in two-phase flow applications. For this case the base functions are only updated along the front. Now, we may formulate the finite dimensional problem. We want to seek $u_{\epsilon,h} \in V_\epsilon^h$ with $u_{\epsilon,h} = \sum_{z \in Z_h^0} p_z \psi_z$ such that

$$\int_{V_z} (\theta(\eta^{u_h}) - \theta^{n-1}) dx - \Delta t \int_{\partial V_z} k_s(x) k_r(\eta^{u_h}) D_x u_{\epsilon,h} \cdot n dl = 0, \tag{4.20}$$

for every control volume $V_z \subset \Omega$. To this equation we can directly apply the linearization procedure described in [17]. Let us here denote

$$r^m = u_{\epsilon,h}^m - u_{\epsilon,h}^{m-1}, \quad m = 1, 2, 3, \dots, \tag{4.21}$$

where $u_{\epsilon,h}^m$ is the iterate of $u_{\epsilon,h}$ at the iteration level m . Thus, we want to find $r^m = \sum_{z \in Z_h^0} r_z^m \psi_z$ such that for $m = 1, 2, 3, \dots$ $\|r^m\| \leq \delta$ with δ being some predetermined error tolerance

$$\int_{V_z} C(\eta^{u_h, m-1}) r^m dx - \Delta t \int_{\partial V_z} k_s(x) k_r(\eta^{u_h, m-1}) D_x r^m \cdot n dl = R^{h, m-1}, \tag{4.22}$$

with

$$R^{h, m-1} = - \int_{V_z} (\theta(\eta^{u_h, m-1}) - \theta^{n-1}) dx + \Delta t \int_{\partial V_z} k_s(x) k_r(\eta^{u_h, m-1}) D_x u_{\epsilon,h}^{m-1} \cdot n dl. \tag{4.23}$$

The superscript m at each of the functions means that the corresponding functions are evaluated at an iteration level m .

We present several numerical experiments that demonstrate the ability of the coarse models presented in the previous subsections. The coarse models are compared with the fine model solved on a fine mesh. We have employed a finite volume difference to solve the fine-scale equations. This solution serves as a reference for the proposed coarse models. The problems that we consider are typical water infiltration into an initially dry soil. The porous medium that we consider is a rectangle of size $L_x \times L_z$

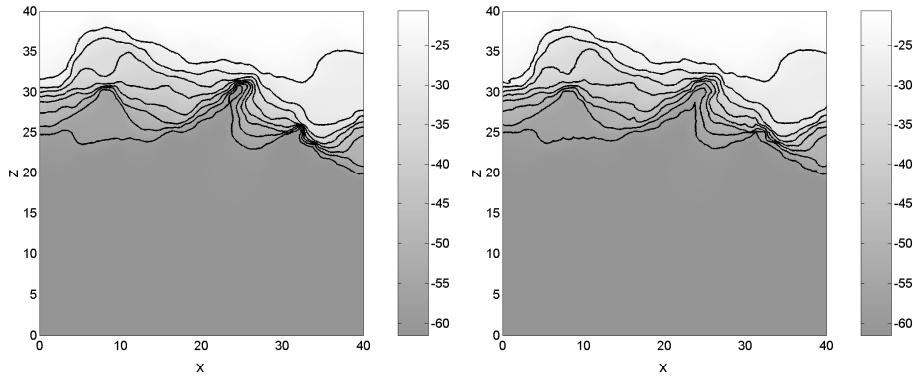


FIG. 4.2. *Haverkamp model with isotropic heterogeneity. Comparison of water pressure between the fine model (left) and the coarse model (right).*

(see Figure 4.1). The fine model uses 256×256 rectangular elements, while the coarse model uses 32×32 rectangular elements.

A realization of the permeability field is generated using geostatistical package GSLIB ([9]). We have used a spherical variogram with prescribed correlation lengths (l_x, l_z) and the variance (σ) for this purpose. All examples use $\sigma = 1.5$.

The first problem is a soil infiltration, which was first analyzed by Haverkamp (cf. [6]). The porous medium dimension is $L_x = 40$ and $L_z = 40$. The boundary conditions are as follows: Γ_L and Γ_R are impermeable, while Dirichlet conditions are imposed on Γ_B and Γ_T , namely $u_T = -21.7$ in Γ_T , and $u_B = -61.5$ in Γ_B . The initial pressure is $u_0 = -61.5$. The constitutive relations use the **Haverkamp model**. The related parameters are as follows: $\alpha = 1.611 \times 10^6$, $\theta_s = 0.287$, $\theta_r = 0.075$, $\beta = 3.96$, $A = 1.175 \times 10^6$, and $\gamma = 4.74$. For this problem we assume that the nonlinearity and heterogeneity are separable, where the latter comes from $K_s(x)$ with $\overline{K_s} = 0.00944$. We assume that appropriate units for these parameters hold. There are two cases considered for this problem, namely, the isotropic heterogeneity with $l_x = l_z = 0.1$, and the anisotropic heterogeneity with $l_x = 0.01$ and $l_z = 0.20$. For the backward Euler scheme, we use $\Delta t = 10$. Note that the large value of Δt is due to the smallness of $\overline{K_s}$ (average magnitude of the diffusion). The comparison is shown in Figures 4.2 and 4.3, where the solutions are plotted at $t = 360$.

The second problem is a soil infiltration through a porous medium whose dimension is $L_x = 1$ and $L_z = 1$. The boundary conditions are as follows: Γ_L and Γ_R are impermeable. Dirichlet conditions are imposed on Γ_B with $u_B = -10$. The boundary Γ_T is divided into three parts. On the middle part, a zero Dirichlet condition is imposed, and the rest are impermeable. The constitutive relations use the **Exponential model** with the following related parameters: $\beta = 0.01$, $\theta_s = 1$, $\overline{K_s} = 1$, and $\overline{\alpha} = 0.01$. The heterogeneity comes from $K_s(x)$ and $\alpha(x)$. Clearly, for this problem the nonlinearity and heterogeneity are not separable. Again, isotropic and anisotropic heterogeneities are considered with $l_x = l_z = 0.1$ and $l_x = 0.20, l_z = 0.01$, respectively. For the backward Euler scheme, we use $\Delta t = 2$. The comparison is shown in Figures 4.4 and 4.5, where the solutions are plotted at $t = 10$.

We note that the problems that we have considered are vertical infiltration on the porous medium. Hence, it is also useful to compare the cross-sectional vertical

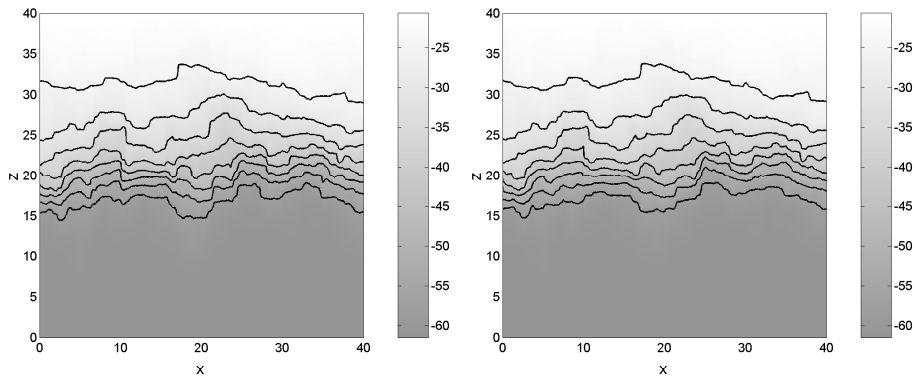


FIG. 4.3. *Haverkamp model with anisotropic heterogeneity. Comparison of water pressure between the fine model (left) and the coarse model (right).*

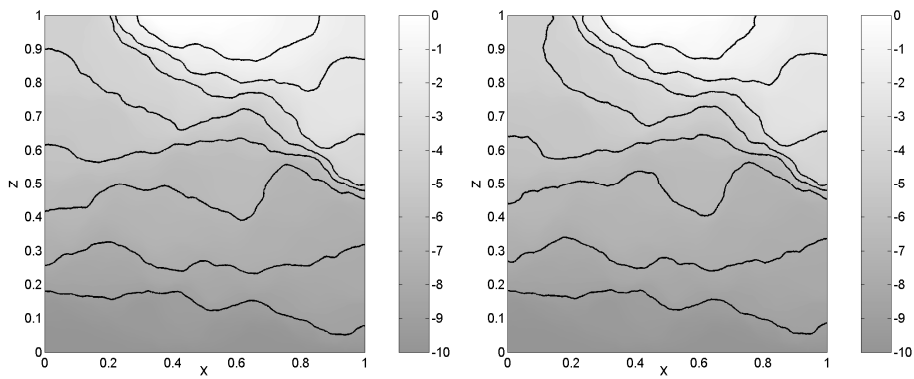


FIG. 4.4. *Exponential model with isotropic heterogeneity. Comparison of water pressure between the fine model (left) and the coarse model (right).*

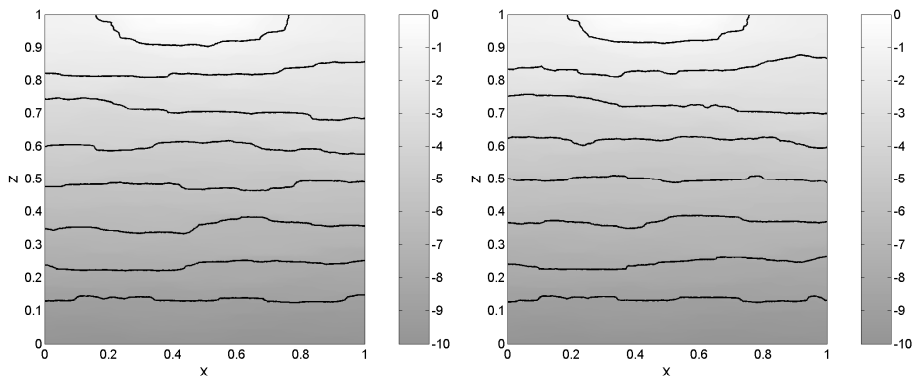


FIG. 4.5. *Exponential model with anisotropic heterogeneity. Comparison of water pressure between the fine model (left) and the coarse model (right).*

velocity that will be plotted against the depth z . Here, the cross-sectional vertical velocity is obtained by taking an average over the horizontal direction (x -axis).

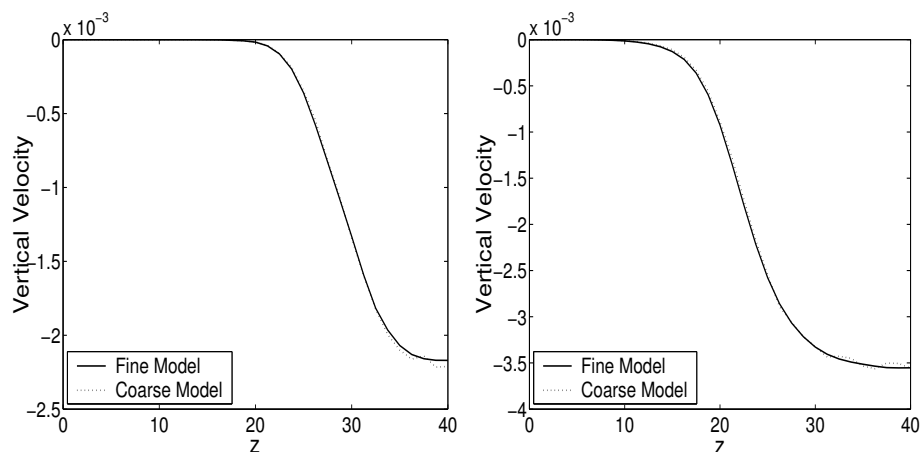


FIG. 4.6. Comparison of vertical velocity on the coarse grid for Haverkamp model: isotropic heterogeneity (left) and anisotropic heterogeneity (right).

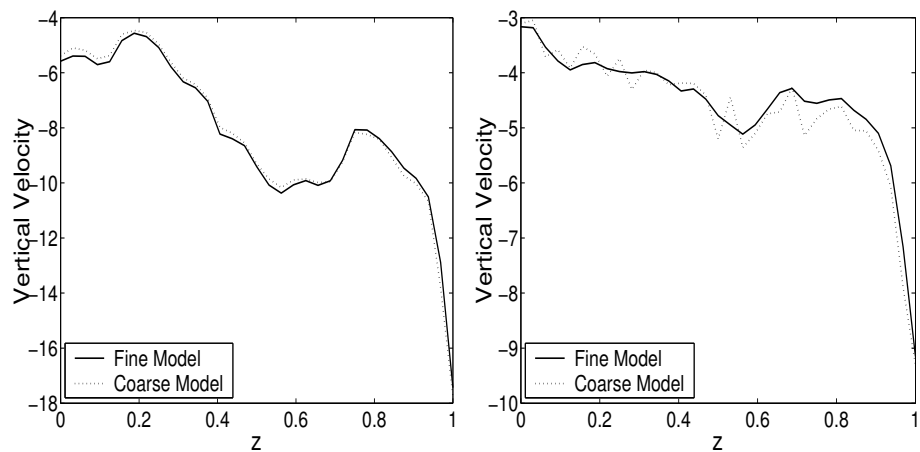


FIG. 4.7. Comparison of vertical velocity on the coarse grid for Exponential model: isotropic heterogeneity (left) and anisotropic heterogeneity (right). The average is taken over the second third of the domain.

Figure 4.6 shows a comparison of the cross-sectional vertical velocity for the Haverkamp model. The average is taken over all the horizontal span because the boundary condition on Γ_T (and also on Γ_B) is all Dirichlet condition. Both plots in this figure show a close agreement between the fine and coarse models. For the Exponential model, as we have described above, there are three different segments for the boundary condition on Γ_T , i.e., a Neumann condition on the first and third part, and a Dirichlet condition on the second/middle part of Γ_T . Thus, we will compare the cross-sectional vertical velocity in each of these segments separately. Figure 4.7

shows the comparison for one of these segments. The agreement between the coarse grid and fine grid calculations is excellent.

5. Extension of MsFEM and concluding remarks

5.1. On the convergence of MsFEM. In this paper we have discussed the convergence of MsFEM in the limit $\epsilon/h \rightarrow 0$ for the periodic problems. This result, we believe, also holds for nonlinear elliptic problems with random homogeneous coefficients, where we assume $a_\epsilon(x, \eta, \xi) = a(T(x/\epsilon)\omega, \eta, \xi)$, where $a(\omega, \eta, \xi)$ is a random homogeneous field. The analysis of the case with random heterogeneities is different from the periodic case and it is currently under investigation. For more general cases without any assumptions on the nature of heterogeneities, one can show (see [13])

$$\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|u_h - u\|_{W_0^{1,p}(\Omega)} = 0, \tag{5.1}$$

(up to a subsequence) where u can be a solution of (1.2) and u_h is a MsFEM solution given by (2.3). The proof of this fact uses only G -convergence results and holds up to a subsequence of ϵ . As we mentioned before, this result can not be improved because there could be infinitely many scales $(\alpha(\epsilon))$ such that $\alpha(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$.

5.2. MsFEM for nonlinear parabolic equations. Consider $u_\epsilon \in W_0$

$$D_t u_\epsilon - \operatorname{div}(a_\epsilon(x, t, u_\epsilon, D_x u_\epsilon)) + a_{0,\epsilon}(x, t, u_\epsilon, D_x u_\epsilon) = f,$$

where $W_0 = \{u \in W, u(x, t = 0) = 0\}$, and

$$\begin{aligned} V_0 &= L^p(0, T, W_0^{1,p}(\Omega)), \\ \bar{V} &= L^p(0, T, W^{1,p}(\Omega)), \\ W &= \{u \in V_0, D_t u \in L^q(0, T, W^{-1,q}(\Omega))\}, \\ \bar{W} &= \{u \in \bar{V}, D_t u \in L^q(0, T, W^{-1,q}(\Omega))\}. \end{aligned} \tag{5.2}$$

Assume $0 = t_0 < t_1 < \dots < t_M = T$, where $\max(t_{i+1} - t_i) = h_t$ and denote $h = \max(h_x, h_t)$. The multiscale mapping for nonlinear parabolic equations, $E^{MsFEM} : S^h \rightarrow V_\epsilon^h$, is constructed in the following way. For each $v_h \in S^h$, $v_{\epsilon,h}(x, t)$ is the solution of

$$D_t v_{\epsilon,h} - \operatorname{div}(a_\epsilon(x, t, \eta^{v_h}, D_x v_{\epsilon,h})) = 0 \text{ in } K \times [t_n, t_{n+1}], \tag{5.3}$$

$v_{\epsilon,h} = v_h$ on ∂K and $v_{\epsilon,h}(t = t_n) = v_h$. Note that E^{MsFEM} is a one-to-one map defined on $\Omega \times [t_n, t_{n+1}]$. MsFEM formulation of the problem is the following. Find $u_h(t) \in S^h$ (and $u_{\epsilon,h} \in V_\epsilon^h$) such that

$$\int_{t_n}^{t_{n+1}} \int_{\Omega} D_t u_h v_h dx dt + \langle A_{\epsilon,h} u_h, v_h \rangle = \int_{t_n}^{t_{n+1}} \int_{\Omega} f v_h dx dt, \tag{5.4}$$

where

$$\langle A_{\epsilon,h} u_h, v_h \rangle = \int_{t_n}^{t_{n+1}} \int_{\Omega} ((a_\epsilon(x, t, \eta^{u_h}, D_x u_{\epsilon,h}), D_x v_h) + a_{0,\epsilon}(x, t, \eta^{u_h}, D_x u_{\epsilon,h}) v_h) dx dt.$$

One can write (5.4) in the following way

$$\int_{\Omega} (u_h(t_{n+1}) - u_h(t_n)) v_h dx + \langle A_{\epsilon,h} u_h, v_h \rangle = \int_{t_n}^{t_{n+1}} \int_{\Omega} f v_h dx dt.$$

Furthermore, taking the present value of u_h , i.e., $u_h(x, t_{n+1})$, in $\langle A_{\epsilon, h} u_h, v_h \rangle$ we obtain the implicit MsFEM scheme and taking the value of u_h at $t = t_n$ we obtain the explicit MsFEM scheme. For the special periodic case, one can use the solution in the period to construct E^{MsFEM} as we did in the elliptic case. Finally, we would like to note that one can develop oversampling techniques for parabolic problems, by extending both temporal and spatial domains of the local problems, and this is a subject of our future research.

Next, we consider some concrete examples.

Example 1. Linear Case.

In the linear case, the multiscale map E^{MsFEM} is linear, and consequently, V_ϵ^h is a linear space. A basis of V_ϵ^h can be found by mapping a basis of S^h , i.e., $V_\epsilon^h = span(\phi_\epsilon^i)$, where ϕ_ϵ^i satisfy

$$D_t \phi_\epsilon^i - \operatorname{div}(a_\epsilon(x, t) D_x \phi_\epsilon^i) = 0 \text{ in } K \times [t_n, t_{n+1}], \tag{5.5}$$

$\phi_\epsilon^i(t = t_n) = \phi_0^i$ on ∂K and $\phi_\epsilon^i = \phi_0^i$, where $S^h = span(\phi_0^i)$. The development of the oversampling techniques for both space and time is currently under investigation.

Example 2. Spatial case

If we assume $a_\epsilon(x, t, \eta, \xi)$ is independent of time, then the following local problems can be solved for the construction of E^{MsFEM} (instead of (5.3))

$$-\operatorname{div}(a_\epsilon(x, \eta^{v_h}, D_x v_{\epsilon, h})) = 0 \text{ in } K \times [t_n, t_{n+1}], \tag{5.6}$$

where $v_{\epsilon, h} = v_h$ on ∂K and $v_{\epsilon, h}(t = t_n) = v_h$. This simplification of the local problem can be understood based on the homogenization of parabolic equations. In particular, the solution of (5.3) can be approximated with the solution of (5.6) in the case of spatial heterogeneities.

5.3. Further generalizations of MsFEM and concluding remarks.

Next, we present the framework of MsFEM for general equations. Consider

$$L_\epsilon u_\epsilon = f, \tag{5.7}$$

where ϵ is a small scale and $L_\epsilon : X \rightarrow Y$ is an operator. Moreover, we assume that L_ϵ G -converges to L^* (up to a sub-sequence), where u is a solution of

$$L^* u = f, \tag{5.8}$$

(we refer to [28], page 14 for the definition of G -convergence for operators). The objective of MsFEM is to approximate u in S^h . Denote S^h a family of finite dimensional space such that it possesses an approximation property (see [37, 29]) as before. Here h is a scale of computation and $h \gg \epsilon$. For (5.7) *multiscale mapping*, $E^{MsFEM} : S^h \rightarrow V_\epsilon^h$, will be defined as follows. For each element $v_h \in S^h$, $v_{\epsilon, h} = E^{MsFEM} v_h$ is defined as

$$L_\epsilon^{map} v_{\epsilon, h} = 0 \text{ in } K, \tag{5.9}$$

where L_ϵ^{map} can be, in general, different from L_ϵ and allows us to capture the effects of the small scales. Moreover, the domains different from the target coarse block K can be used in the computations of the local solutions. To solve (5.9) one needs to impose boundary and initial conditions. This issue needs to be resolved on a case by case basis, and the main idea is to interpolate v_h onto the underlying fine grid. Further,

we seek a solution of (5.7) in V_ϵ^h as follows. Find $u_h \in S^h$ (consequently $u_{\epsilon,h} \in V_\epsilon^h$) such that

$$\langle L_\epsilon^{global} u_{\epsilon,h}, v_h \rangle = \langle f, v_h \rangle, \quad \forall v_h \in S^h, \tag{5.10}$$

where $\langle u, v \rangle$ denotes the duality between X and Y , and L_ϵ^{global} can be, in general, different from L_ϵ . For example, for nonlinear elliptic equations we have $L_\epsilon u = -\text{div}(a_\epsilon(x, u, D_x u)) + a_{0,\epsilon}(x, u, D_x u)$, $L_\epsilon^{map} u = \text{div}(a_\epsilon(x, \eta^u, D_x u))$ in K , and $L_\epsilon^{global} = \text{div}(a_\epsilon(x, \eta^u, D_x u)) + a_{0,\epsilon}(x, \eta^u, D_x u)$ in K . The convergence of MsFEM is to show that $u_h \rightarrow u$ and $u_{\epsilon,h} \rightarrow u_\epsilon$, where $u_{\epsilon,h} = E^{MsFEM} u_h$ in appropriate space. The correct choices of L_ϵ^{map} and L_ϵ^{global} are the essential part of MsFEM and guarantees the convergence of the method.

In conclusion, we have presented a natural extension of MsFEM to nonlinear problems. This is accomplished by considering a multiscale map instead of the base functions that are considered in linear MsFEM [19]. Our approaches share some common elements with recently introduced HMM [10], where macroscopic and microscopic solvers are also needed. In general, the finding of “correct” macroscopic and microscopic solvers is the main difficulty of the multiscale methods. Our approaches follow MsFEM and, consequently, finite element methods constitute its main ingredient. The resonance errors, that arise in linear problems also arise in nonlinear problems. Note that the resonance errors are the common feature of multiscale methods unless periodic problems are considered and the solutions of the local problems in an exact period are used. To reduce the resonance errors we use an oversampling technique and show that the error can be greatly reduced by sampling from the larger domains. The multiscale map for MsFEM uses the solutions of the local problems in the target coarse block. This way one can sample the heterogeneities of the coarse block. If there is a scale separation and, in addition, some kind of periodicity, one can use the solutions of the smaller size problems to approximate the multiscale map. Note that a potential disadvantage of periodicity assumption is that the periodicity can act to disrupt large-scale connectivity features of the flow. For the examples similar to the non-periodic ones considered in this paper, with the use of the smaller size problems for approximating the solutions of the local problems, we have found very large errors (of order 50 percent).

Appendix A. The proof of Lemma 3.3.

Let $\tilde{v}_{\epsilon,h} = v_{\epsilon,h} - v_h$, where $v_{\epsilon,h} = E^{MsFEM} v_h$. It follows that $\tilde{v}_{\epsilon,h} \in W_0^{1,p}(K)$ satisfies the following problem:

$$-\text{div}(a_\epsilon(x, \eta^{v_h}, D_x \tilde{v}_{\epsilon,h} + D_x v_h)) = 0 \text{ in } K. \tag{A.1}$$

Using (A.1), applying Green’s Theorem we have the following estimate:

$$\begin{aligned} & \langle A_{\epsilon,h} v_h, v_h \rangle \\ &= \sum_{K \in \mathbf{K}^h} \int_K (a_\epsilon(x, \eta^{v_h}, D_x v_h + D_x \tilde{v}_{\epsilon,h}), D_x v_h + D_x \tilde{v}_{\epsilon,h}) dx \\ &+ \sum_{K \in \mathbf{K}^h} \int_K a_{0,\epsilon}(x, \eta^{v_h}, D_x v_{\epsilon,h}) v_h dx \end{aligned}$$

$$\begin{aligned}
 &= \sum_{K \in \mathbf{K}^h} \int_K [(a_\epsilon(x, \eta^{v_h}, D_x v_h + D_x \tilde{v}_{\epsilon,h}), D_x v_h + D_x \tilde{v}_{\epsilon,h}) + a_{0,\epsilon}(x, \eta^{v_h}, D_x v_{\epsilon,h}) \eta^{v_h}] dx \\
 &\quad + \sum_{K \in \mathbf{K}^h} \int_K a_{0,\epsilon}(x, \eta^{v_h}, D_x v_{\epsilon,h}) (v_h - \eta^{v_h}) dx. \tag{A.2}
 \end{aligned}$$

Further, using the coercivity condition (3.3) and the fact that $|v_h - \eta^{v_h}| \leq Ch|D_x v_h|$ in each K (note that $D_x v_h$ is constant in each K) we have

$$\begin{aligned}
 \langle A_{\epsilon,h} v_h, v_h \rangle &\geq c \sum_{K \in \mathbf{K}^h} \int_K |D_x v_h + D_x \tilde{v}_{\epsilon,h}|^p dx - c_1 h \sum_{K \in \mathbf{K}^h} \int_K a_{0,\epsilon}(x, \eta^{v_h}, D_x v_{\epsilon,h}) D_x v_h dx \\
 &\geq (c - c_1 h) \sum_{K \in \mathbf{K}^h} \int_K |D_x v_{\epsilon,h}|^p dx. \tag{A.3}
 \end{aligned}$$

Next, we discuss some special cases, where the coercivity can be easily shown. If $p=2$,

$$\begin{aligned}
 &\sum_{K \in \mathbf{K}^h} \int_K |D_x v_{\epsilon,h}|^2 dx \\
 &= \sum_{K \in \mathbf{K}^h} \int_K |D_x v_h + D_x \tilde{v}_{\epsilon,h}|^2 dx \\
 &= \sum_{K \in \mathbf{K}^h} \int_K |D_x v_h|^2 dx + \sum_{K \in \mathbf{K}^h} \int_K (D_x v_h, D_x v_{\epsilon,h}) dx + \sum_{K \in \mathbf{K}^h} \int_K |D_x v_{\epsilon,h}|^2 dx \\
 &= \sum_{K \in \mathbf{K}^h} \int_K |D_x v_h|^2 dx + \sum_{K \in \mathbf{K}^h} \int_K |D_x v_{\epsilon,h}|^2 dx \\
 &\geq \sum_{K \in \mathbf{K}^h} \int_K |D_x v_h|^2 dx. \tag{A.4}
 \end{aligned}$$

Here we have used the fact that $\tilde{v}_{\epsilon,h} = 0$ on ∂K , and $D_x v_h$ is constant in each K . One can also easily show the coercivity, using rescaling arguments, if $a_\epsilon(x, \eta, \xi) = a_\epsilon(x, \xi)$ and $a_\epsilon(x, \lambda \xi) = |\lambda|^{p-1} b_\epsilon(x, \xi)$.

Next we analyze the general case. Denote by K_r a reference triangle (with the size of order one) and introduce the change of variables $z = x/h$. Using Trace Theorem (see e.g., [27], page 30) $\|u\|_{L^p(\partial Q)} \leq c \|u\|_{W^{1,p}(Q)}$ and the fact $\|D_x u\|_{L^p(Q)} \geq c \|u - f(u)\|_{W^{1,p}(Q)}$ (see e.g., [15], page 490) where Q is a domain with a Lipschitz boundary and $f(u)$ can be taken to be the average of u on ∂Q we get the following.

$$\begin{aligned}
 &\sum_{K \in \mathbf{K}^h} \int_K |D_x v_{\epsilon,h}|^p dx \\
 &= \sum_{K \in \mathbf{K}^h} h^{d-p} \int_{K_r} |D_z v_{\epsilon,h}|^p dz \geq \sum_{K \in \mathbf{K}^h} h^{d-p} \|v_{\epsilon,h} - \tilde{\eta}^{v_h}\|_{W^{1,p}(K_r)} \\
 &\geq c \sum_{K \in \mathbf{K}^h} h^{d-p} \int_{\partial K_r} |v_h - \tilde{\eta}^{v_h}|^p dz_l \\
 &= c \sum_{K \in \mathbf{K}^h} h^{d-p} \int_{\partial K_r} |(D_z v_h, z - z_0)|^p dz_l \\
 &= c \sum_{K \in \mathbf{K}^h} h^d |D_x v_h|^p \int_{\partial K_r} |(e_{D_z v_h}, z - z_0)|^p dz_l. \tag{A.5}
 \end{aligned}$$

Here $\tilde{\eta}^{v_h}$ is the average of v_h along the boundaries, $\tilde{\eta}^{v_h} = \frac{1}{|\partial K_r|} \int_{\partial K_r} v_h dz_l$ and $v_h = \tilde{\eta}^{v_h} + (D_z v_h, z - z_0)$, where $z_0 = \frac{1}{|\partial K_r|} \int_{\partial K_r} z dz_l$. Denote

$$C(e_{D_z v_h}) = \int_{\partial K_r} |(e_{D_z v_h}, z - z_0)|^p dz_l.$$

To finish the proof, we need only to claim that $C(e_\xi)$ is bounded from below independent of ξ . By contradiction suppose the claim is not true. Then there exists a sequence $\{e_{\xi_n}\}$ that has a sub-sequence (denoted by the same notation) such that $e_{\xi_n} \rightarrow e_*$ and $C(e_{\xi_n}) \rightarrow 0$ as $n \rightarrow \infty$. Because $C(e_\xi)$ is continuous it follows that $C(e_*) = 0$. This further implies that $(e_*, z) = 0$ on ∂K_r , and hence $e_* = 0$. This is a contradiction. Finally, denoting the lower bound of $C(e_{D_x v_h})$ by c_0 , we have

$$\langle A_{\epsilon, h} v_h, v_h \rangle \geq c \sum_{K \in \mathbf{K}^h} h^d |D_x v_h|^p C(e_{D_x v_h}) \geq c c_0 \sum_K \int_K |D_x v_h|^p dx = c \|D_x v_h\|_{p, \Omega}^p. \tag{A.6}$$

Appendix B. Proof of Lemma 3.4.

First we show the following fact.

PROPOSITION B.1. *Let $v_\epsilon - v_0 \in W_0^{1,p}(K)$ satisfies the following problem:*

$$-\operatorname{div}(a_\epsilon(x, \eta, D_x v_\epsilon)) = 0 \text{ in } K \tag{B.1}$$

where η is constant in K . Then

$$\|D_x v_\epsilon\|_{p, K} \leq c(|K|^{\frac{1}{p}} + \|\eta\|_{p, K} + \|D_x v_0\|_{p, K}). \tag{B.2}$$

Proof. Let $\tilde{v}_\epsilon = v_\epsilon - v_0$. It follows that \tilde{v}_ϵ satisfies the following problem:

$$-\operatorname{div}(a_\epsilon(x, \eta, D_x(\tilde{v}_\epsilon + v_0))) = 0 \text{ in } K \quad \text{and} \quad \tilde{v}_\epsilon = 0 \text{ on } \partial K. \tag{B.3}$$

Multiplying (B.3) with v_ϵ , applying Green’s Theorem, and using the fact that $\tilde{v}_\epsilon = 0$ on ∂K , we immediately obtain the following equality:

$$\int_K (a_\epsilon(x, \eta, D_x v_\epsilon), D_x v_\epsilon) dx = \int_K (a_\epsilon(x, \eta, D_x v_\epsilon), D_x v_0) dx. \tag{B.4}$$

Next we use coercivity and polynomial growth properties of $a_\epsilon(x, \eta, \xi)$ to bound (B.4) from below and above, respectively. Thus by applying Holder’s and Young’s inequalities we have

$$\begin{aligned} c_2 \|D_x v_\epsilon\|_{p, K}^p &\leq c_1 \int_K (1 + |\eta|^{p-1} + |D_x v_\epsilon|^{p-1}) |D_x v_0| dx \\ &\leq c_1 \left(\int_K (1 + |\eta|^p + |D_x v_\epsilon|^p) dx \right)^{\frac{1}{q}} \|D_x v_0\|_{p, K} \\ &\leq \frac{c_1 \delta}{q} \int_K (1 + |\eta|^p + |D_x v_\epsilon|^p) dx + \frac{c_1}{p \delta} \|D_x v_0\|_{p, K}^p. \end{aligned} \tag{B.5}$$

The claim in this proposition is obtained from this inequality by choosing $\delta > 0$ appropriately. \square

Next we prove Lemma 3.4. Let $\tilde{v}_\epsilon = v_\epsilon - v_0$ and $\tilde{w}_\epsilon = w_\epsilon - w_0$. It follows that \tilde{v}_ϵ and \tilde{w}_ϵ satisfy the following problems respectively:

$$-\operatorname{div}(a_\epsilon(x, \eta, D_x(\tilde{v}_\epsilon + v_0))) = 0 \text{ in } K \quad \text{and} \quad \tilde{v}_\epsilon = 0 \text{ on } \partial K, \tag{B.6}$$

$$-\operatorname{div}(a_\epsilon(x, \eta, D_x(\tilde{w}_\epsilon + w_0))) = 0 \text{ in } K \quad \text{and} \quad \tilde{w}_\epsilon = 0 \text{ on } \partial K. \tag{B.7}$$

Using the monotonicity property of $a_\epsilon(x, \eta, \xi)$ and applying Green’s Theorem along with the fact that $\tilde{v}_\epsilon = \tilde{w}_\epsilon = 0$ on ∂K , we immediately obtain the following inequality:

$$\begin{aligned} & c_1 \|D_x(v_\epsilon - w_\epsilon)\|_{p,K}^p \\ &= c_1 \|D_x(\tilde{v}_\epsilon + v_0) - D_x(\tilde{w}_\epsilon + w_0)\|_{p,K}^p \\ &\leq \int_K (a_\epsilon(x, \eta, D_x v_\epsilon) - a_\epsilon(x, \eta, D_x w_\epsilon), D_x(v_0 - w_0)) \, dx \\ &\leq c_4 \int_K H(\eta, D_x v_\epsilon, \eta, D_x w_\epsilon, p-1-s) |D_x(v_\epsilon - w_\epsilon)|^s |D_x(v_0 - w_0)| \, dx, \end{aligned} \tag{B.8}$$

where on the last line we have used the continuity property of $a_\epsilon(x, \eta, \xi)$ and H is defined in (3.4). Applying Holder’s and Young’s inequalities appropriately we have

$$\begin{aligned} & \|D_x(v_\epsilon - w_\epsilon)\|_{p,K}^p \\ &\leq c \left(\int_K H(\eta, D_x v_\epsilon, \eta, D_x w_\epsilon, p) \, dx \right)^{\frac{p-s-1}{p}} \|D_x(v_0 - w_0)\|_{p,K} \|D_x(v_\epsilon - w_\epsilon)\|_{p,K}^s \\ &\leq c \frac{\delta s}{p} \|D_x(v_\epsilon - w_\epsilon)\|_{p,K}^p + c \frac{p-s}{\delta p} \left(\int_K H(\eta, D_x v_\epsilon, \eta, D_x w_\epsilon, p) \, dx \right)^{\frac{p-s-1}{p-s}} \\ &\quad \|D_x(v_0 - w_0)\|_{p,K}^{\frac{p}{p-s}}. \end{aligned} \tag{B.9}$$

Applying Proposition B.1 and choosing $\delta > 0$ appropriately, we obtain the desired estimate.

Appendix C. Proof Proposition 3.6.

By change of variables, it is sufficient to show that

$$\|P_{\eta,\xi}\|_{p,Y}^p \leq c(1 + |\eta|^p + |\xi|^p), \tag{C.1}$$

where Y is the unit square. Applying monotonicity and polynomial growth properties of $a(y, \eta, \xi)$ we have

$$\begin{aligned} \|P_{\eta,\xi}\|_{p,Y}^p &= \int_Y |P_{\eta,\xi} - 0|^p \, dy \\ &\leq c \int_Y (a(y, \eta, P_{\eta,\xi}) - a(y, \eta, 0), P_{\eta,\xi}) \, dy \\ &= c \int_Y (a(y, \eta, P_{\eta,\xi}), \xi) \, dy - c \int_Y (a(y, \eta, 0), P_{\eta,\xi}) \, dy \\ &\leq c \int_Y (1 + |\eta|^{p-1} + |P_{\eta,\xi}|^{p-1}) |\xi| \, dy + c \int_Y (1 + |\eta|^{p-1}) |P_{\eta,\xi}| \, dy. \end{aligned} \tag{C.2}$$

Next we use Holder’s inequality with $r_1 = p/(p-1)$ and $r_2 = p$ on both terms and afterward apply Young’s inequality, so that for some $\beta > 0$ we have

$$\|P_{\eta,\xi}\|_{p,Y}^p \leq c_1(\beta) (1 + |\eta|^p + |\xi|^p) + c_2(\beta) \|P_{\eta,\xi}\|_{p,Y}^p. \tag{C.3}$$

Here $c_2(\beta) \rightarrow 0$ as $\beta \rightarrow 0$. Choosing β appropriately, we obtain the desired estimate.

Appendix D. Proof Lemma 3.8.

The following lemma will be used in the proof (see (D.12)).

LEMMA D.1. *If $u_k \rightarrow 0$ in $L^r(\Omega)$ ($1 < r < \infty$) as $k \rightarrow \infty$ then*

$$\int_{\Omega} \nu(u_k) |v_k|^p dx \rightarrow 0, \text{ as } k \rightarrow \infty$$

for all v_k either (1) compact in $L^p(\Omega)$ or (2) uniformly bounded in $L^{p+\alpha}(\Omega)$, $\alpha > 0$. Here $\nu(t)$ is continuity modulus defined previously and $1 < p < \infty$.

Proof. Because u_k converges in $L^r(\Omega)$ it converges in measure. Consequently, for any $\delta > 0$ there exists Ω_δ and k_0 such that $meas(\Omega \setminus \Omega_\delta) < \delta$ and $\nu(u_k) < \delta$ in Ω_δ for $k > k_0$. Thus

$$\int_{\Omega} \nu(u_k) |v_k|^p dx = \int_{\Omega_\delta} \nu(u_k) |v_k|^p dx + \int_{\Omega \setminus \Omega_\delta} \nu(u_k) |v_k|^p dx \leq C\delta + C \int_{\Omega \setminus \Omega_\delta} |v_k|^p dx. \tag{D.1}$$

Next we use the fact that if (1) or (2) is satisfied then the set v_k has equi-absolute continuous norm [23] (i.e., for any $\epsilon > 0$ there exists $\zeta > 0$ such that $meas(\Omega_\zeta) < \zeta$ implies $\|P_{\Omega_\zeta} v_k\|_p < \epsilon$, where $P_D f = \{f(x), \text{ if } x \in D; 0 \text{ otherwise}\}$). Consequently, the second term on the right side of (D.1) converges to zero as $\delta \rightarrow 0$. \square

Given $v_h \in S^h$, we set the corrector \mathcal{P} as in (3.18). By adding and subtracting terms, we have the following equality:

$$\langle A_{\epsilon,h} v_h - A^* v_h, w_h \rangle = \sum_{K \in \mathbf{K}^h} (I_K + II_K + III_K) + \sum_{K \in \mathbf{K}^h} (i_K + ii_K + iii_K), \tag{D.2}$$

where

$$\begin{aligned} I_K &= \int_K (a(x/\epsilon, \eta^{v_h}, D_x v_{\epsilon,h}) - a(x/\epsilon, \eta^{v_h}, \mathcal{P}), D_x w_h) dx, \\ II_K &= \int_K (a(x/\epsilon, \eta^{v_h}, \mathcal{P}) - a^*(\eta^{v_h}, D_x v_h), D_x w_h) dx, \\ III_K &= \int_K (a^*(\eta^{v_h}, D_x v_h) - a^*(v_h, D_x v_h), D_x w_h) dx, \end{aligned} \tag{D.3}$$

and

$$\begin{aligned} i_K &= \int_K (a_0(x/\epsilon, \eta^{v_h}, D_x v_{\epsilon,h}) - a_0(x/\epsilon, \eta^{v_h}, \mathcal{P})) w_h dx, \\ ii_K &= \int_K (a_0(x/\epsilon, \eta^{v_h}, \mathcal{P}) - a_0^*(\eta^{v_h}, D_x v_h)) w_h dx, \\ iii_K &= \int_K (a_0^*(\eta^{v_h}, D_x v_h) - a_0^*(v_h, D_x v_h)) w_h dx. \end{aligned} \tag{D.4}$$

We will show the convergence corresponding to higher order terms, I_K, II_K, III_K . The convergence estimates for i_K, ii_K, iii_K are the same and can be obtained in a very analogous manner.

Step 1: estimate of I_K

Using continuity property (3.5) and Holder’s inequality, I_K is estimated in the following way:

$$\begin{aligned}
 I_K &\leq c \int_K |D_x v_{\epsilon,h} - \mathcal{P}|^s H(\eta^{v_h}, D_x v_{\epsilon,h}, \eta^{v_h}, \mathcal{P}, p-1-s) |D_x w_h| dx \\
 &\leq c \|D_x v_{\epsilon,h} - \mathcal{P}\|_{p,K}^s \left(\int_K H(\eta^{v_h}, D_x v_{\epsilon,h}, \eta^{v_h}, \mathcal{P}, p) dx \right)^{\frac{p-1-s}{p}} \|D_x w_h\|_{p,K}, \tag{D.5}
 \end{aligned}$$

where H is defined by (3.4). Note that on the second term in the last line of (D.5) we have

$$\begin{aligned}
 &\int_K H(\eta^{v_h}, D_x v_{\epsilon,h}, \eta^{v_h}, \mathcal{P}, p) dx \\
 &= 1 + 2 \|\eta^{v_h}\|_{p,K}^p + \|D_x v_{\epsilon,h}\|_{p,K}^p + \|\mathcal{P}\|_{p,K}^p \\
 &\leq c \left(1 + \|\eta^{v_h}\|_{p,K}^p + \|D_x v_{\epsilon,h}\|_{p,K}^p + \|D_x v_h\|_{p,K}^p + \|\epsilon D_x N_{\eta^{v_h}, D_x v_h}\|_{p,K}^p \right). \tag{D.6}
 \end{aligned}$$

Using Proposition B.1 and a technique similar to the one in the proof of Lemma 3.5, (see (3.27)) we have

$$\int_K H(\eta^{v_h}, D_x v_{\epsilon,h}, \eta^{v_h}, \mathcal{P}, p) dx \leq c \left(|K| + \|v_h\|_{p,K}^p + \|D_x v_h\|_{p,K}^p \right). \tag{D.7}$$

With this estimate along with (3.19) it follows from (D.5) that

$$\begin{aligned}
 \sum_{K \in \mathbf{K}^h} I_K &\leq c \|D_x v_{\epsilon,h} - \mathcal{P}\|_{p,\Omega}^s \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p \right)^{\frac{p-1-s}{p}} \|D_x w_h\|_{p,\Omega} \\
 &\leq c \left(\frac{\epsilon}{h} \right)^{\frac{s}{p(p-s)}} \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p \right)^{\frac{s}{p}} \\
 &\quad \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p \right)^{\frac{p-1-s}{p}} \|D_x w_h\|_{p,\Omega} \\
 &\leq c \left(\frac{\epsilon}{h} \right)^{\frac{s}{p(p-s)}} \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p \right)^{\frac{1}{q}} \|D_x w_h\|_{p,\Omega}. \tag{D.8}
 \end{aligned}$$

By Lemma 3.5 the last inequality vanishes as ϵ approaching zero.

Step 2: estimate of II_K

Let $I_\epsilon^K = \{i \in \mathbb{Z}^d : Y_\epsilon^i \subset K\}$ and $J_\epsilon^K = \{i \in \mathbb{Z}^d : Y_\epsilon^i \cap K \neq \emptyset, K \setminus Y_\epsilon^i \neq \emptyset\}$. Let $E_\epsilon^K = \cup_{i \in I_\epsilon^K} Y_\epsilon^i$ and $F_\epsilon^K = \cup_{i \in J_\epsilon^K} Y_\epsilon^i$. Then we may break up the integration II_K into the sum of integral over E_ϵ^K and $K \setminus E_\epsilon^K$. By (3.16) and the fact that $D_x w_h$ is constant in K , we have the following estimate:

$$\begin{aligned}
 II_K &= \sum_{i \in I_\epsilon^K} \int_{Y_\epsilon^i} (a(x/\epsilon, \eta^{v_h}, \mathcal{P}) - a^*(\eta^{v_h}, D_x v_h), D_x w_h) dx \\
 &\quad + \int_{K \setminus E_\epsilon^K} (a(x/\epsilon, \eta^{v_h}, \mathcal{P}) - a^*(\eta^{v_h}, D_x v_h), D_x w_h) dx \\
 &\leq \int_{F_\epsilon^K} |(a(x/\epsilon, \eta^{v_h}, \mathcal{P}) - a^*(\eta^{v_h}, D_x v_h), D_x w_h)| dx. \tag{D.9}
 \end{aligned}$$

It follows by applying Holder’s inequality appropriately and using Proposition 3.6 that

$$\begin{aligned}
 & \sum_{K \in \mathbf{K}^h} II_K \\
 \leq & \sum_{K \in \mathbf{K}^h} \sum_{i \in J_\epsilon^K} \int_{Y_\epsilon^i} (|(a(x/\epsilon, \eta^{v_h}, \mathcal{P}), D_x w_h)| + |(a^*(\eta^{v_h}, D_x v_h), D_x w_h)|) dx \\
 \leq & c \sum_{K \in \mathbf{K}^h} \sum_{i \in J_\epsilon^K} \int_{Y_\epsilon^i} H(\eta^{v_h}, \mathcal{P}, \eta^{v_h}, D_x v_h, p-1) |D_x w_h| dx \\
 \leq & c \left(\sum_{K \in \mathbf{K}^h} \sum_{i \in J_\epsilon^K} \int_{Y_\epsilon^i} H(\eta^{v_h}, \mathcal{P}, \eta^{v_h}, D_x v_h, p) dx \right)^{\frac{1}{q}} \left(\sum_{K \in \mathbf{K}^h} \sum_{i \in J_\epsilon^K} \int_{Y_\epsilon^i} |D_x w_h|^p dx \right)^{\frac{1}{p}} \\
 \leq & c \left(\sum_{K \in \mathbf{K}^h} \sum_{i \in J_\epsilon^K} (1 + |\eta^{v_h}|^p + |D_x v_h|^p) |Y_\epsilon^i| \right)^{\frac{1}{q}} \left(\sum_{K \in \mathbf{K}^h} \sum_{i \in J_\epsilon^K} |D_x w_h|^p |Y_\epsilon^i| \right)^{\frac{1}{p}} \\
 \leq & c \left(\sum_{K \in \mathbf{K}^h} |K| (1 + |\eta^{v_h}|^p + |D_x v_h|^p) \frac{|F_\epsilon^K|}{|K|} \right)^{\frac{1}{q}} \left(\sum_{K \in \mathbf{K}^h} |K| |D_x w_h|^p \frac{|F_\epsilon^K|}{|K|} \right)^{\frac{1}{p}} \\
 \leq & c \max_K \left(\frac{|F_\epsilon^K|}{|K|} \right)^{\frac{1}{q}} \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p \right)^{\frac{1}{q}} \max_K \left(\frac{|F_\epsilon^K|}{|K|} \right)^{\frac{1}{q}} \left(\|D_x w_h\|_{p,\Omega}^p \right)^{\frac{1}{p}} \\
 \leq & c \left(\frac{\epsilon}{h} \right) \left(|\Omega| + \|v_h\|_{p,\Omega}^p + \|D_x v_h\|_{p,\Omega}^p \right)^{\frac{1}{q}} \cdot \|D_x w_h\|_{p,\Omega} \tag{D.10}
 \end{aligned}$$

This expression vanishes as ϵ approaches to zero.

Step 3: estimate of III_K

Using (3.5) for the homogenized fluxes and Holder’s inequality we estimate III_K in the following way:

$$\begin{aligned}
 III_K & \leq c \int_K H(\eta^{v_h}, D_x v_h, v_h, D_x v_h, p-1) \nu(|\eta^{v_h} - v_h|) |D_x w_h| dx \\
 & \leq c \left(\int_K H(\eta^{v_h}, D_x v_h, v_h, D_x v_h, p) \nu(|\eta^{v_h} - v_h|)^q dx \right)^{\frac{1}{q}} \|D_x w_h\|_{p,K}. \tag{D.11}
 \end{aligned}$$

It follows that

$$\begin{aligned}
 \sum_{K \in \mathbf{K}^h} III_K & \leq c \left(\int_\Omega H(\eta^{v_h}, D_x v_h, v_h, D_x v_h, p) \nu(|\eta^{v_h} - v_h|)^q dx \right)^{\frac{1}{q}} \|D_x w_h\|_{p,\Omega} \\
 & \leq c \left(\int_\Omega (|v_h|^p + |D_x v_h|^p) \nu(|\eta^{v_h} - v_h|)^q dx \right)^{\frac{1}{q}} \|D_x w_h\|_{p,\Omega}. \tag{D.12}
 \end{aligned}$$

Because $D_x v_h \in L^{p+\alpha}(\Omega)$, $D_x w_h \in L^p(\Omega)$ and $\eta^{v_h} - v_h$ converges to zero in $L^p(\Omega)$ it follows from Lemma D.1 that $\sum_{K \in \mathbf{K}^h} III_K$ vanishes as $\epsilon \rightarrow 0$.

Acknowledgments. The research of Y. E. is partially supported by NSF grants DMS-0327713 and EIA-0218229. The research of T.Y.H. is partially supported by the NSF ITR grant ACI-0204932. We would like to acknowledge anonymous reviewers for their helpful comments which helped to improve the quality of the paper.

REFERENCES

- [1] T. Arbogast, *An overview of subgrid upscaling for elliptic problems in mixed form*, in Current Trends in Scientific Computing (Xi'an, 2002), Contemp. Math., Amer. Math. Soc., Providence, RI, 329, 21–32, 2003.
- [2] T. Arbogast, *Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems*, SIAM J. Numer. Anal., 42, 2, 576–598, 2004.
- [3] I. Babuška, G. Caloz and 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.
- [4] I. Babuška and E. Osborn, *Generalized finite element methods: their performance and their relation to mixed methods*, SIAM J. Numer. Anal., 20, 510–536, 1983.
- [5] F. Brezzi, *Interacting with the subgrid world*, in Numerical analysis 1999 (Dundee), Chapman & Hall/CRC, Boca Raton, FL, 69–82, 2000.
- [6] M. A. Celia, E. T. Bouloutas and R. L. Zarba, *A general mass-conservative numerical solution for the unsaturated flow equation*, Water Resour. Res., 26, 7, 1483–1496, 1990.
- [7] S.- S. Chow, *Finite element error estimates for nonlinear elliptic equations of monotone type*, Numer. Math., 54, 373–393, 1989.
- [8] G. Dal Maso and A. Defranceschi, *Correctors for the homogenization of monotone operators*, Differential Integral Equations, 3, 1151–1166, 1990.
- [9] C. V. Deutsch and A. G. Journel, *GSLIB: Geostatistical Software Library and User's Guide, 2nd Edition*, Oxford University Press, New York, 1998.
- [10] W. E and B. Engquist, *The heterogeneous multi-scale methods*, Comm. Math. Sci., 1, 1, 87–132, 2003.
- [11] Y. Efendiev and A. Pankov, *Homogenization of nonlinear random parabolic operators*, submitted to EJDE (available at <http://www.math.tamu.edu/~yalchin.efendiev/ep-hom-parab.ps>).
- [12] Y. Efendiev and A. Pankov, *Meyers type estimates for approximate solutions of nonlinear elliptic equations and their applications*, submitted to Num. Math., (available at <http://www.math.tamu.edu/~yalchin.efendiev/ep-meyers-elliptic.ps>).
- [13] Y. Efendiev and A. Pankov, *Numerical homogenization of nonlinear random parabolic operators*, SIAM Multiscale Model. and Simul., 2, 2, 237–268, 2004.
- [14] Y. R. Efendiev, T. Y. Hou and X. H. Wu, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Num. Anal., 37, 888–910, 2000.
- [15] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*, Applied Mathematical Sciences, Springer-Verlag, New York, 159, 2004.
- [16] R. Eymard, T. Gallouët and R. Herbin, *Finite volume methods*, in Handbook of Numerical Analysis, Vol. VII, Handb. Numer. Anal., VII, North-Holland, Amsterdam, 713–1020, 2000.
- [17] V. Ginting, *Computational Upscaled Modeling of Heterogeneous Porous Media Flows Utilizing Finite Volume Method*, PhD thesis, Texas A&M University, College Station, 2004.
- [18] T. Hou, X. Wu and Y. Zhang, *Removing the cell resonance error in the multiscale finite element method via a petrov-galerkin formulation*, Comm. in Math. Sci., 2, 2, 185–205, 2004.
- [19] T. Y. Hou and X. H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, Journal of Computational Physics, 134, 169–189, 1997.
- [20] 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.
- [21] T. Hughes, G. Feijoo, L. Mazzei and J. Quincy, *The variational multiscale method - a paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg, 166, 3–24, 1998.
- [22] V. V. Jikov, S. M. Kozlov and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, New York, 1994.
- [23] M. A. Krasnosel'skiĭ, P. P. Zabreĭko, E. I. Pustyl'nik and P. E. Sobolevskiĭ, *Integral Operators In Spaces of Summable Functions*, Noordhoff International Publishing, Leiden, 1976. (Translated from the Russian by T. Ando, Monographs and Textbooks on Mechanics of Solids and Fluids, Mechanics: Analysis.)
- [24] J. Lions, *Quelques Methodes de Resoluton des Problemes aux Limites Non Lineaires*, Dunod, Paris, 1969.

- [25] A.-M. Matache and C. Schwab, *Homogenization via p -FEM for problems with microstructure*, in Proceedings of the Fourth International Conference on Spectral and High Order Methods (ICOSAHOM 1998) (Herzliya), 33, 43-59, 2000.
- [26] N. G. Meyers and A. Elcrat, *Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions*, Duke Math. J., 42, 121-136, 1975.
- [27] J. Nečas, *Introduction to the Theory of Nonlinear Elliptic Equations*, A Wiley-Interscience Publication, John Wiley & Sons Ltd., Chichester, 1986. Reprint of the 1983 edition.
- [28] A. Pankov, *G-convergence and Homogenization of Nonlinear Partial Differential Operators*, Kluwer Academic Publishers, Dordrecht, 1997.
- [29] W. V. Petryshyn, *On the approximation-solvability of equations involving A -proper and pseudo- A -proper mappings*, Bull. Amer. Math. Soc., 81, 223-312, 1975.
- [30] L. Richards, *Capillary conduction of liquids through porous mediums*, Physics, 318-333, 1931.
- [31] G. Sangalli, *Capturing small scales in elliptic problems using a residual-free bubbles finite element method*, SIAM Multiscale Model. Simul., 1, 485-503, 2003 (electronic).
- [32] R. E. Showalter, *Monotone operators in Banach space and nonlinear partial differential equations*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 49, 1997.
- [33] I. V. Skrypnik, *Methods for analysis of nonlinear elliptic boundary value problems*, Translations of Mathematical Monographs, American Mathematical Society, Providence, RI, 139, 1994. (Translated from the 1990 Russian original by Dan D. Pascali).
- [34] M. T. van Genuchten, *A closed-form equation for predicting the hydraulic conductivity of unsaturated soils*, Soil. Sci. Soc. Am. J., 44, 892-898, 1980.
- [35] A. W. Warrick, *Time-dependent linearized infiltration: III. strip and disc sources*, Soil. Sci. Soc. Am. J., 40, 639-643, 1976.
- [36] X. Wen, L. Durlofsky and M. Edwards, *Use of border regions for improved permeability upscaling*, Mathematical Geology, 35, 521-547, 2003.
- [37] E. Zeidler, *Nonlinear Functional Analysis and Its Applications. II/B*, Springer-Verlag, New York, 1990. (Nonlinear monotone operators, Translated from the German by the author and Leo F. Boron).