

## NUMERICAL HOMOGENIZATION AND CORRECTORS FOR NONLINEAR ELLIPTIC EQUATIONS\*

Y. EFENDIEV<sup>†</sup> AND A. PANKOV<sup>‡</sup>

**Abstract.** In this paper we consider numerical homogenization and correctors for nonlinear elliptic equations. The numerical correctors are constructed for operators with homogeneous random coefficients. The construction employs two scales, one a physical scale and the other a numerical scale. A numerical homogenization technique is proposed and analyzed. This procedure is developed within finite element formulation. The convergence of the numerical procedure is presented for the case of general heterogeneities using  $G$ -convergence theory. The proposed numerical homogenization procedure for elliptic equations can be considered as a generalization of multiscale finite element methods to nonlinear equations. Using corrector results we construct an approximation of oscillatory solutions. Numerical examples are presented.

**Key words.** homogenization, multiscale, upscaling, scale-up, random, nonlinear, elliptic, finite element

**AMS subject classification.** 65N99

**DOI.** 10.1137/S0036139903424886

**1. Introduction.** Consider the nonlinear elliptic equations

$$(1) \quad -\operatorname{div}(a_\epsilon(x, u_\epsilon, Du_\epsilon)) + a_{0,\epsilon}(x, u_\epsilon, Du_\epsilon) = f, \quad u_\epsilon \in W_0^{1,p}(Q).$$

Here  $\epsilon$  denotes the small scale of the problem. Direct numerical simulations of these kinds of problems are difficult because of scale disparity. Our objective is to find the approximation of the homogenized solution without solving the fine scale problem; i.e., (1) is solved on a grid of size  $h$ , where  $h \gg \epsilon$ . The numerical procedure introduced for this purpose can be regarded as numerical homogenization. The numerical homogenization procedure for (1) should account for the functional dependence of the macroscopic quantities on the solution and its gradients. Our motivation in considering (1) mostly stems from the applications of flow in porous media (multiphase flow in saturated porous media and flow in unsaturated porous media) and enhanced diffusion due to nonlinear heterogeneous convection, though many applications of nonlinear elliptic equations of these kinds occur in transport problems.

In this paper we consider two issues: (1) the calculation of the correctors and (2) the computation of the homogenized solution. The homogenization of nonlinear elliptic equations in a random media has been studied previously (see, e.g., [17]). It was shown that a solution  $u_\epsilon$  converges (up to a subsequence) to  $u$  in an appropriate norm and where  $u \in W_0^{1,p}(Q)$  is a solution of

$$(2) \quad -\operatorname{div}(a^*(x, u, Du)) + a_0^*(x, u, Du) = f.$$

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

---

\*Received by the editors March 23, 2003; accepted for publication (in revised form) January 29, 2004; published electronically September 24, 2004.

<http://www.siam.org/journals/siap/65-1/42488.html>

<sup>†</sup>Department of Mathematics, Texas A&M University, College Station, TX 77843-3368 (efendiev@math.tamu.edu). The work of this author is partially supported by NSF grants DMS 0327713 and EIA-0218229.

<sup>‡</sup>Department of Mathematics, College of William & Mary, Williamsburg, VA 23187-8795 (pankov@math.wm.edu).

strictly stationary fields with respect to spatial variables. In these cases one has an auxiliary problem of calculating  $a^*$  and  $a_0^*$ . The numerical homogenization procedure presented in this paper does not use the auxiliary problem for the calculation of the approximation of homogenized solutions.

To construct the numerical correctors we use two scales, a physical scale and a numerical scale that is much larger than the physical one, and construct the correctors in each numerical coarse block. The convergence for the corrector is obtained. These results show us a way to obtain numerically the fine scale features of the solution. We would like to note that the computation of the oscillation of solutions is important for the application to flow in porous media and other transport problems.

We present a procedure for calculating a coarse solution, the solution at the length scales  $h$ ,  $1 \gg h \gg \epsilon$ . Our numerical homogenization procedure is based on general finite element computations of the coarse scale equations. It selectively solves the required local problems that reduce overall computations even in the periodic case. The solutions of the local problems are uniquely determined, which makes our discrete operator single-valued. The convergence of the numerical method is presented for general kinds of heterogeneities using  $G$ -convergence theory. Moreover, we show that the numerical homogenization approach presented in this paper can be considered as a generalization of multiscale finite element methods introduced in [10]. A related work in multiscale computations involves generalized finite element methods [2], residual free bubbles [3, 19], the variational multiscale method [12], two-scale finite element methods [15], two-scale conservative subgrid approaches [1], and the heterogeneous multiscale method (HMM) [6].

Some numerical examples are considered in this paper. We study numerically the effect of enhanced diffusion due to heterogeneous nonlinear convection,

$$\frac{\partial u_\epsilon}{\partial t} + \frac{1}{\epsilon} v_\epsilon(x) \cdot DF(u_\epsilon) - d\Delta u_\epsilon = f.$$

Since the elliptic part does not depend on  $t$ , the theory developed previously can be applied. In this application we are interested in the effect of the enhanced diffusion due to heterogeneous nonlinear convection. More precisely, assuming the existence of homogeneous stream function for the velocity field and zero mean drift, we calculate the approximation of the enhanced diffusion due to the convection using Buckley–Leverett flux that describes the convection. Other numerical examples for Richards equations are also studied.

The paper is organized as follows. In the next section we present some basic facts that are used later in the analysis. Section 3 is devoted to the construction of a numerical corrector and its convergence. Section 4 is devoted to the calculation of the homogenized solution and its analysis. In section 5 we present numerical results.

**2. Preliminaries.** We start with a description of random homogeneous fields on  $R^d$ , which are shown to be useful in homogenization problems (see, e.g., [13]). Let  $(\Omega, \Sigma, \mu)$  be a probability space. A random homogeneous field is a measurable function on  $\Omega$  and  $f(T(x)\omega)$  are realizations of the random field. The realizations are well-defined measurable functions on  $R^d$  for almost all  $\omega \in \Omega$ . Consider a  $d$  dimensional dynamical system on  $\Omega$ ,  $T(x) : \Omega \rightarrow \Omega$ ,  $x \in R^d$ , that satisfies the following conditions: (1)  $T(0) = I$ , and  $T(x+y) = T(x)T(y)$ ; (2)  $T(x) : \Omega \rightarrow \Omega$  preserve the measure  $\mu$  on  $\Omega$ ; (3) for any measurable function  $f(\omega)$  on  $\Omega$ , the function  $f(T(x)\omega)$  defined on  $R^d \times \Omega$  is also measurable (see [13, 18]). Let  $L^p(\Omega)$  denote the space of all  $p$ -integrable

functions on  $\Omega$ . Then  $U(x)f(\omega) = f(T(x)\omega)$  defines a  $d$ -parameter group of isometries in the space  $L^p(\Omega)$ , and  $U(x)$  is strongly continuous [13, 17]. Further, we assume that the dynamical system  $T$  is ergodic; i.e., any measurable  $T$ -invariant function on  $\Omega$  is constant. Denote by  $\langle \cdot \rangle$  the mean value over  $\Omega$ ,

$$\langle f \rangle = \int_{\Omega} f(\omega) d\mu(\omega) = E(f).$$

Now we explain briefly the relation between the standard definition of random homogeneous fields and the one we introduced here following, e.g., [17]. Let  $\Xi$  be a probability space endowed with a probability measure  $P$ . Let  $f$  be a random vector valued function, i.e., a measurable map  $f : \Xi \times R^d \rightarrow R^N$ .  $f$  is a random homogeneous field if all its finite dimensional distributions are translation invariant. The latter means that for any  $x^1, x^2, \dots, x^k \in R^d$ , and any Borel subsets  $B_1, B_2, \dots, B_k \subset R^N$ ,

$$P\{\xi \in \Xi : f(\xi, x^1 + h) \in B_1, \dots, f(\xi, x^k + h) \in B_k\}$$

is independent of  $h \in R^d$ . Consider a new probability space  $\Omega$  and a dynamical system  $T(x)$  acting on  $\Omega$ . We define  $\Omega$  to be the set of all measurable functions  $\omega : R^d \rightarrow R^N$  and set  $T(x)\omega(y) = \omega(x + y)$ ,  $x, y \in R^d$ . Let  $\mathcal{F}$  be the  $\sigma$ -algebra generated by ‘‘cylinder’’ sets, i.e., the sets of the form  $B = \{\omega : \omega(x^1) \in B_1, \dots, \omega(x^k) \in B_k\}$ , where  $x^1, x^2, \dots, x^k \in R^d$  and  $B_1, B_2, \dots, B_k$  are Borel subsets in  $R^N$ . We define the measure  $\mu$  on ‘‘cylinder’’ sets by  $\mu(B) = P\{\xi \in \Xi : f(\xi, \cdot) \in B\}$  and then extend it to  $\mathcal{F}$  by  $\sigma$ -additivity. Thus, the probability space  $\Omega$  and the measure-preserving dynamical system  $T(x)$ ,  $x \in R^d$ , on  $\Omega$  are constructed. Moreover, consider the  $\mu$ -measurable function  $\hat{f} : \Omega \rightarrow R^N$  defined by the formula  $\hat{f}(\omega) = \omega(0)$ . Then  $f(\xi, x) = \hat{f}(T(x)\omega)$ , where  $\omega(\cdot) = f(\xi, \cdot)$ . More examples regarding the construction of  $T$  can be found in [13].

Denote by  $\partial_{\omega}^i$  the generator of  $U(x)$  along the  $i$ th coordinate direction, i.e.,

$$\partial_{\omega}^i = \lim_{\delta \rightarrow 0} \frac{f(T(\delta e_i)\omega) - f(\omega)}{\delta}.$$

The domains  $D^i$  of  $\partial_{\omega}^i$  are dense in  $L^2(\Omega)$ , and the intersection of all  $D^i$  is also dense.

Next, following [17], we define potential and solenoidal fields. A vector field  $f \in L^p(\Omega)$  is said to be potential (or solenoidal, respectively) if its generic realization  $f(T(x)\omega)$  is a *potential* (or *solenoidal*, respectively) vector field in  $R^d$ . Denote by  $L_{pot}^p(\Omega)$  (respectively,  $L_{sol}^p(\Omega)$ ) the subspace of  $L^p(\Omega)$  that consists of all potential (respectively, solenoidal) vector fields. Introduce the following notation:

$$V_{pot}^p = \{f \in L_{pot}^p(\Omega), \langle f \rangle = 0\}, \quad V_{sol}^p = \{f \in L_{sol}^p(\Omega), \langle f \rangle = 0\}.$$

The following properties are known (see [17, page 138]):

$$L_{pot}^p(\Omega) = V_{pot}^p \oplus R^d, \quad L_{sol}^p(\Omega) = V_{sol}^p \oplus R^d, \quad L_{sol}^q(\Omega) = (V_{pot}^p)^{\perp}, \quad L_{pot}^q(\Omega) = (V_{sol}^p)^{\perp}.$$

Consider  $u_{\epsilon} \in W_0^{1,p}(Q)$ ,

$$(3) \quad -\operatorname{div}(a(T(x/\epsilon)\omega, u_{\epsilon}, Du_{\epsilon})) + a_0(T(x/\epsilon)\omega, u_{\epsilon}, Du_{\epsilon}) = f \quad \text{in } Q,$$

where  $f$  is a deterministic function that does not depend on  $\epsilon$  and is sufficiently smooth, and  $Q \subset R^d$  is a domain with Lipschitz boundaries.

Assume for all  $\omega \in \Omega$

$$(4) \quad (a(\omega, \eta, \xi_1) - a(\omega, \eta, \xi_2), \xi_1 - \xi_2) \geq C(1 + |\xi_1| + |\xi_2|)^{p-\beta} |\xi_1 - \xi_2|^\beta,$$

$$(5) \quad |a(\omega, \eta, \xi)| + |a_0(\omega, \eta, \xi)| \leq C(1 + |\eta| + |\xi|)^{p-1},$$

$$(6) \quad \begin{aligned} & |a(\omega, \eta_1, \xi_1) - a(\omega, \eta_2, \xi_2)| + |a_0(\omega, \eta_1, \xi_1) - a_0(\omega, \eta_2, \xi_2)| \\ & \leq C(1 + |\eta_1|^{p-1} + |\eta_2|^{p-1} + |\xi_1|^{p-1} + |\xi_2|^{p-1})\nu(|\eta_1 - \eta_2|) \\ & \quad + C(1 + |\eta_1|^{p-1-s} + |\eta_2|^{p-1-s} + |\xi_1|^{p-1-s} + |\xi_2|^{p-1-s})|\xi_1 - \xi_2|^s, \end{aligned}$$

where  $0 < s \leq 1$ ,  $\beta \geq \max(p, 2)$ ,  $p > 1$ . Here  $\nu(r)$  is a continuity modulus; i.e.,  $\nu(r)$  is a nondecreasing continuous function on  $[0, +\infty)$  such that  $\nu(0) = 0$ ,  $\nu(r) > 0$  if  $r > 0$  and  $\nu(r) = 1$  if  $r > 1$ , and  $\nu(r_1 + r_2) \leq C(\nu(r_1) + \nu(r_2))$ . For the existence of the solution we need a coercivity condition,

$$(7) \quad (a(\omega, \eta, \xi), \xi) + a_0(\omega, \eta, \xi)\eta \geq C|\xi|^p - C_1.$$

It is known (e.g., [17]) that, as  $\epsilon \rightarrow 0$ ,  $Du_\epsilon$  converges to  $Du$  weakly in  $L^p(Q)^d$  for almost every  $\omega$ , and  $u$  is the solution of

$$(8) \quad -\operatorname{div}(a^*(u, Du)) + a_0^*(u, Du) = f, \quad u \in W_0^{1,p}(Q).$$

Further,  $a^*$  and  $a_0^*$  can be constructed using the solution of the following auxiliary problem. Given  $\eta \in R$  and  $\xi \in R^d$ , define  $w_{\eta,\xi} \in V_{pot}^p$  such that

$$(9) \quad a(\omega, \eta, \xi + w_{\eta,\xi}(\omega)) \in L_{sol}^q(\Omega)^d.$$

Then  $a^*(\eta, \xi)$  and  $a_0(\eta, \xi)$  are defined as

$$(10) \quad \begin{aligned} a^*(\eta, \xi) &= \langle a(\omega, \eta, \xi + w_{\eta,\xi}(\omega)) \rangle, \\ a_0^*(\eta, \xi) &= \langle a_0(\omega, \eta, \xi + w_{\eta,\xi}(\omega)) \rangle. \end{aligned}$$

Moreover,  $a^*(\eta, \xi)$  and  $a_0^*(\eta, \xi)$  satisfy estimates similar to those of  $a$  and  $a_0$  with different constants [17].

*Remark 2.1.* We would like to note that  $G$ -convergence and homogenization results presented in [17] were formulated under weaker than (4) conditions. In particular, it is assumed that

$$(11) \quad (a(\omega, \eta, \xi_1) - a(\omega, \eta, \xi_2), \xi_1 - \xi_2) \geq C(1 + |\eta| + |\xi_1| + |\xi_2|)^{p-\beta} |\xi_1 - \xi_2|^\beta.$$

It turns out that  $G$ -convergence and homogenization results hold under more general assumptions such as (4). The proof is identical to the one presented in [17]. Moreover, following [17], it can be easily shown that the homogenized operator is also coercive and satisfies (7).

Throughout the paper  $C$  denotes a generic constant,  $\|\cdot\|_p$  denotes  $L^p(Q)$  (or the broken norm), and  $L^p(Q)^d$  norms and  $q$  are defined by  $1/p + 1/q = 1$ . The notation a.e. (almost every) is often omitted.

**3. Two-scale correctors.** The corrector results obtained in this section will be used in the approximation of solution gradients. The importance of this approximation is motivated by some applications in which details of the fluxes play a key role in a physical phenomenon (e.g., flow in porous media). For the construction we assume that the homogenized solution is computed with a reliable accuracy in an appropriate norm which will be specified later. In the next section we will propose a numerical procedure for the computation of the homogenized solution for more general heterogeneities. For the construction of the correctors we introduce two scale correctors, where one scale represents the numerical scale  $h$  and the other the physical scale  $\epsilon$ .

Define  $M_h\phi(x)$  in the following way:

$$M_h\phi(x) = \sum_i 1_{Q_i} \frac{1}{|Q_i|} \int_{Q_i} \phi(y) dy,$$

where  $\bigcup Q_i = Q$ . Here  $Q_i$  are domains with diameter of order  $h$ , e.g., finite element triangles or some unions of the triangles. Note that  $M_h\phi \rightarrow \phi$  in  $L^p(Q)$  as  $h \rightarrow 0$  (see [22]). Further, define

$$(12) \quad P(T(y)\omega, \eta, \xi) = \xi + w_{\eta, \xi}(T(y)\omega),$$

where  $w_{\eta, \xi} \in V_{pot}^p(\Omega)$  is the solution of the auxiliary problem  $a(\omega, \eta, \xi + w_{\eta, \xi}(\omega)) \in L_{sol}^q(\Omega)^d$ . Here  $w_{\eta, \xi}(T(y)\omega)$  satisfies

$$-\operatorname{div}(a(T(y)\omega, \eta, \xi + w_{\eta, \xi}(T(y)\omega))) = 0$$

in the sense of distribution [17, p. 155].

The main result of this section regarding the convergence of the correctors is the following.

**THEOREM 3.1.** *Let  $u_\epsilon$  and  $u$  be solutions of (3) and (8), respectively, and let  $P$  be defined by (12) in each  $Q_i$ . Then*

$$(13) \quad \lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \int_Q |P(T(x/\epsilon)\omega, M_h u, M_h Du) - Du_\epsilon|^p dx = 0$$

*$\mu$ -a.e.*

We will omit  $\mu$ -a.e. notation in further analysis. To make the expressions in the proof more concise we introduce the notation

$$\mathcal{P}_\epsilon = P(T(x/\epsilon)\omega, M_h u, M_h Du).$$

Theorem 3.1 indicates that the gradient of solutions can be approximated by  $P(T(x/\epsilon)\omega, M_h u, M_h Du)$ . This quantity can be computed based on  $M_h Du$  and  $M_h u$ , i.e., the gradient of the coarse solution in each coarse block, as we will show later. The following lemma [4] will be used in the proof.

**LEMMA 3.2.** *For any  $\phi_1$  and  $\phi_2$  belonging to  $L_p(Q)$  we have*

$$(14) \quad \|\phi_1 - \phi_2\|_{p, Q} \leq C \left( \int_Q |\phi_1 - \phi_2|^\beta (1 + |\phi_1| + |\phi_2|)^{p-\beta} dx \right)^{1/\beta} \\ \times (|Q|^{1/p} + \|\phi_1\|_{p, Q} + \|\phi_2\|_{p, Q})^{(\beta-p)/\beta}.$$

For the proof of Theorem 3.1 we need the following lemma.

LEMMA 3.3. For every  $\eta \in R$  and  $\xi \in R^d$

$$\|P(\omega, \eta, \xi)\|_{p, \Omega}^p \leq C(1 + |\eta|^p + |\xi|^p).$$

*Proof.* Using Lemma 3.2 and (4), we obtain

$$\begin{aligned} \|P\|_{p, \Omega}^p &\leq C(1 + \|P\|_{p, \Omega}^p)^{(\beta-p)/\beta} \left( \int_{\Omega} |P|^\beta (1 + |P|)^{p-\beta} d\mu(\omega) \right)^{p/\beta} \\ &\leq C\delta^{(p-\beta)/p} \int_{\Omega} |P|^\beta (1 + |P|)^{p-\beta} d\mu(\omega) + C\delta(1 + \|P\|_{p, \Omega}^p). \end{aligned}$$

With a suitable choice of  $\delta$  and using (4) and (5), we get

$$\begin{aligned} \|P\|_{p, \Omega}^p &\leq C + C \int_{\Omega} |P|^\beta (1 + |P|)^{p-\beta} d\mu(\omega) \leq C + C \int_{\Omega} (a(\omega, \eta, P) - a(\omega, \eta, 0), P) d\mu(\omega) \\ &\leq C + C \left| \int_{\Omega} (a(\omega, \eta, P), P) d\mu(\omega) \right| + \left| \int_{\Omega} (a(\omega, \eta, 0), P) d\mu(\omega) \right| \\ &\leq C + \left| \int_{\Omega} (a(\omega, \eta, P), \xi) d\mu(\omega) \right| + (1 + |\eta|^{p-1}) \left| \int_{\Omega} P d\mu(\omega) \right| \\ &\leq C + C\delta_1 \|P\|_{p, \Omega}^p + C\delta_1^{-1/(p-1)} |\eta|^p + C \int_{\Omega} (1 + |\eta| + |P|)^{p-1} |\xi| d\mu(\omega) \\ &\leq C\delta_2 \|P\|_{p, \Omega}^p + C\delta_2^{-1/(p-1)} (1 + |\xi|^p) + C + C\delta_1 (|\eta|^p + \|P\|_{p, \Omega}^p) + C\delta_1^{-1/(p-1)} |\eta|^p. \end{aligned}$$

With an appropriate choice of  $\delta_1$  and  $\delta_2$ , we obtain the desired result.  $\square$

It follows from Lemma 3.3 that  $P(T(y)\omega, \eta, \xi) \in L_{loc}^p(R^d)^d$  for a.e.  $\omega$  and for each  $\eta \in R$ ,  $\xi \in R^d$ . The next lemma will be also used in the proof of Theorem 3.1.

LEMMA 3.4. If  $u_k \rightarrow 0$  in  $L^r(Q)$  ( $1 < r < \infty$ ) as  $k \rightarrow \infty$ , then

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

for all  $v_k$  either (1) compact in  $L^p(Q)$  or (2) bounded in  $L^{p+\alpha}(Q)$ ,  $\alpha > 0$ . Here  $\nu(r)$  is the continuity modulus defined previously (see (6)) and  $1 < p < \infty$ .

*Proof.* Since  $u_k$  converges in  $L^r$ , it converges in measure. Consequently, for any  $\delta > 0$  there exists  $Q_\delta$  and  $k_0$  such that  $\text{meas}(Q \setminus Q_\delta) < \delta$  and  $\nu(u_k) < \delta$  in  $Q_\delta$  for  $k > k_0$ . Thus

$$\begin{aligned} (15) \quad \int_Q \nu(u_k) |v_k|^p dx &= \int_{Q_\delta} \nu(u_k) |v_k|^p dx + \int_{Q \setminus Q_\delta} \nu(u_k) |v_k|^p dx \\ &\leq C\delta + C \int_{Q \setminus Q_\delta} |v_k|^p dx. \end{aligned}$$

Next we use the fact that if (1) or (2) is satisfied, then the set  $v_k$  has equi-absolute continuous norm [14] (i.e., for any  $\epsilon > 0$  there exists  $\zeta > 0$  such that  $\text{meas}(Q_\zeta) < \zeta$  implies  $\|P_{Q_\zeta} v_k\|_{p, Q} < \epsilon$ , where  $P_D f = \{f(x) \text{ if } x \in D; 0 \text{ otherwise}\}$ ). Consequently, the second term on the right-hand side (r.h.s.) of (15) converges to zero as  $\delta \rightarrow 0$ .  $\square$

The proof of Theorem 3.1 will be based on the following estimate:

$$\begin{aligned} (16) \quad \int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^p dx &\leq \int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^\beta (1 + |\mathcal{P}_\epsilon| + |Du_\epsilon|)^{p-\beta} dx (|Q|)^{1/p} \\ &\quad + \|\mathcal{P}_\epsilon\|_{p, Q} + \|Du_\epsilon\|_{p, Q}^{(\beta-p)/p}. \end{aligned}$$

$\|Du_\epsilon\|_{p,Q}$  is uniformly bounded for a.e.  $\omega$ .  $\|\mathcal{P}_\epsilon\|_{p,Q}$  is also uniformly bounded since  $M_h u$  and  $M_h Du$  are bounded in  $L^p(Q)$  and  $L^p(Q)^d$ , respectively. Thus it remains to estimate  $\int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^\beta (1 + |\mathcal{P}_\epsilon| + |Du_\epsilon|)^{p-\beta} dx$ . For this term, using (4), we have

$$\begin{aligned}
 & \int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^\beta (1 + |\mathcal{P}_\epsilon| + |Du_\epsilon|)^{p-\beta} dx \\
 (17) \quad & \leq C \left| \int_Q (a(T(x/\epsilon)\omega, u_\epsilon, \mathcal{P}_\epsilon) - a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), \mathcal{P}_\epsilon - Du_\epsilon) dx \right| \\
 & \leq C \left| \int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon) - a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), \mathcal{P}_\epsilon - Du_\epsilon) dx \right| \\
 & \quad + C \left| \int_Q (a(T(x/\epsilon)\omega, u_\epsilon, \mathcal{P}_\epsilon) - a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), \mathcal{P}_\epsilon - Du_\epsilon) dx \right|.
 \end{aligned}$$

To prove Theorem 3.1 we will need to estimate the first and second terms on the r.h.s. of (17). For the first term we have

$$\begin{aligned}
 (18) \quad & C \int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon) - a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), \mathcal{P}_\epsilon - Du_\epsilon) dx \\
 & = C \int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), \mathcal{P}_\epsilon) dx - C \int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), Du_\epsilon) dx \\
 & \quad - C \int_Q (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), \mathcal{P}_\epsilon) dx + C \int_Q (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), Du_\epsilon) dx.
 \end{aligned}$$

We will investigate the r.h.s. of (18) in the limit as  $\epsilon \rightarrow 0$ . For the first term of the r.h.s. of (18) we have the following convergence.

LEMMA 3.5.

$$\int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), \mathcal{P}_\epsilon) dx \rightarrow \int_Q (a^*(M_h u, M_h Du), M_h Du) dx$$

as  $\epsilon \rightarrow 0$ .

*Proof.*

$$\begin{aligned}
 & \int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), \mathcal{P}_\epsilon) dx \\
 & = \sum_i \int_{Q_i} (a(T(x/\epsilon)\omega, \eta_i, \xi_i + w_{\eta_i, \xi_i}(T(x/\epsilon)\omega)), \xi_i + w_{\eta_i, \xi_i}(T(x/\epsilon)\omega)) dx \\
 & = \sum_i \int_{Q_i} (a(T(x/\epsilon)\omega, \eta_i, \xi_i + w_{\eta_i, \xi_i}(T(x/\epsilon)\omega)), \xi_i) dx \\
 & \quad + \sum_i \int_{Q_i} (a(T(x/\epsilon)\omega, \eta_i, \xi_i + w_{\eta_i, \xi_i}(T(x/\epsilon)\omega)), w_{\eta_i, \xi_i}(T(x/\epsilon)\omega)) dx \\
 & \rightarrow \sum_i \int_{Q_i} 1_{Q_i}(a^*(\eta_i, \xi_i), \xi_i) dx
 \end{aligned}$$

as  $\epsilon \rightarrow 0$ . In the last step we have used the Birkhoff ergodic theorem (see [13]) as well as (9), (10), and  $w_{\eta, \xi} \in V_{pot}^p$ . Next we note that the limit can be written as

$$\sum_i \int_{Q_i} 1_{Q_i}(a^*(\xi_i), \xi_i) dx = \int_Q (a^*(M_h u, M_h Du), M_h Du) dx. \quad \square$$

For the second term of the r.h.s. of (18) we have the following convergence.

LEMMA 3.6.

$$\int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), Du_\epsilon) dx \rightarrow \int_Q (a^*(M_h u, M_h Du), Du) dx$$

as  $\epsilon \rightarrow 0$ .

*Proof.*

$$\int_Q (a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), Du_\epsilon) dx = \sum_i \int_{Q_i} (a(T(x/\epsilon)\omega, \eta_i P(T(x/\epsilon)\omega, \eta_i, \xi_i)), Du_\epsilon) dx.$$

$Du_\epsilon$  is bounded in  $L^p(Q)^d$  for a.e.  $\omega$ . To show that  $a(T(x/\epsilon)\omega, P(T(x/\epsilon)\omega, \eta_i, \xi_i))$  is bounded in  $L^r(Q_i)^d$ , where  $r > q$ , we will use Meyers' theorem [16]. Since  $-\operatorname{div}(a(T(x/\epsilon)\omega, \eta_i, P(T(x/\epsilon)\omega, \eta_i, \xi_i))) = 0$  in  $3 \times Q_i$  (where  $3 \times Q_i$  is a domain that contains  $Q_i$  and is surrounded with a ring of size  $Q_i$ ), using Meyers' theorem we can conclude that

$$\|P(T(x/\epsilon)\omega, \eta_i, \xi_i)\|_{p+\eta, Q_i} \leq C \|P(T(x/\epsilon)\omega, \eta_i, \xi_i)\|_{p, 3 \times Q_i},$$

where  $C$  is independent of  $\omega$  and depends only on operator constants. Note that  $P \in L^p_{loc}(R^d)^d$ . Since  $\|P(T(x/\epsilon)\omega, \eta_i, \xi_i)\|_{p, 3 \times Q_i}$  is bounded for a.e.  $\omega$  (see Lemma 3.3),  $\|P(T(x/\epsilon)\omega, \eta_i, \xi_i)\|_{p+\alpha, Q_i}$  is also bounded for a.e.  $\omega$ . From here, using bounds for  $a(T(y)\omega, \eta, \xi)$ , we can easily obtain that  $a(T(x/\epsilon)\omega, \eta_i, P(T(x/\epsilon)\omega, \eta_i, \xi_i))$  is bounded in  $L^r(Q_i)^d$ , where  $r > q$  for a.e.  $\omega$ . Indeed,

$$\begin{aligned} & \int_{Q_i} |a(T(x/\epsilon)\omega, \eta_i, P(T(x/\epsilon)\omega, \eta_i, \xi_i)) - a(T(x/\epsilon)\omega, \eta_i, 0)|^r dx \\ & \leq C \int_{Q_i} (|1 + \eta_i + P(T(x/\epsilon)\omega, \eta_i, \xi_i)|)^{(p-2)r} |P(T(x/\epsilon)\omega, \eta_i, \xi_i)|^r dx \\ & \leq C (\|P\|_{r, Q_i} + \|P\|_{(p-1)r, Q_i}). \end{aligned}$$

Since  $P$  is in  $L^{p+\alpha}(Q_i)^d$  for a.e.  $\omega$ , we can pick  $r = q + \alpha/(p-1)$ . Consequently,  $(a(T(x/\epsilon)\omega, \eta_i, \xi_i + w_{\eta_i, \xi_i}(T(x/\epsilon)\omega)), Du_\epsilon)$  is bounded in  $L^\sigma(Q_i)^d$ ,  $\sigma > 1$ , for every  $\eta_i$  and  $\xi_i$ . Thus it contains a subsequence that weak\* converges to  $g_i$  for any  $i$  and a.e.  $\omega$ . Since  $-\operatorname{div}(a(T(x/\epsilon)\omega, \eta_i, P(T(x/\epsilon)\omega, \eta_i, \xi_i))) = 0$  in  $Q_i$ , using a compensated compactness argument we can obtain that as  $\epsilon \rightarrow 0$ ,  $g_i = (a^*(\eta_i, \xi_i), Du)$ . The latter is true because  $Du_\epsilon$  weakly converges to  $Du$  in  $L^p(Q)^d$  for a.e.  $\omega$  and  $a(T(x/\epsilon)\omega, \eta_i, P(T(x/\epsilon)\omega, \eta_i, \xi_i))$  weakly converges to  $a^*(\eta_i, \xi_i)$  in  $L^r(Q)$ . The fact that  $Du_\epsilon$  weakly converges to  $Du$  for a.e.  $\omega$  follows from general  $G$ -convergence results [17], and the weak convergence of  $a(T(x/\epsilon)\omega, \eta_i, P(T(x/\epsilon)\omega, \eta_i, \xi_i))$  is a consequence of the Birkhoff ergodic theorem. Consequently,

$$\begin{aligned} & \sum_i \int_{Q_i} (a(T(x/\epsilon)\omega, \eta_i, P(T(x/\epsilon)\omega, \eta_i, \xi_i)), Du_\epsilon) dx \\ & \rightarrow \sum_i \int_{Q_i} (a^*(\eta_i, \xi_i), Du) dx = \int_Q (a^*(M_h u, M_h Du), Du) dx. \quad \square \end{aligned}$$

For the third term of the r.h.s. of (18) we have the following convergence.

LEMMA 3.7.

$$\int_Q (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), \mathcal{P}_\epsilon) dx \rightarrow \int_Q (a^*(u, Du), M_h Du) dx$$

as  $\epsilon \rightarrow 0$ .



*Proof.*

$$\int_Q (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), \mathcal{P}_\epsilon) dx = \sum_i \int_{Q_i} (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), P(T(x/\epsilon)\omega, \eta_i, \xi_i)) dx.$$

Since  $|a(\omega, \eta, \xi)| \leq C(1 + |\eta|^{p-1} + |\xi|^{p-1})$  and  $P(T(x/\epsilon)\omega, \eta_i, \xi_i)$  converges to  $\xi_i$  in  $L^p(Q)^d$  and is bounded in  $L^{p+\eta}(Q)^d$ , similar to the analysis for the Lemma 3.6 we can obtain that

$$\begin{aligned} & \sum_i \int_{Q_i} (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), P(T(x/\epsilon)\omega, \eta_i, \xi_i)) dx \\ & \rightarrow \sum_i \int_{Q_i} (a^*(u, Du), \xi_i) dx = \int_Q (a^*(u, Du), M_h Du) dx. \quad \square \end{aligned}$$

For the fourth term of the r.h.s. of (18), we have the following.

LEMMA 3.8.

$$\int_Q (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), Du_\epsilon) dx \rightarrow \int_Q (a^*(u, Du), Du) dx$$

as  $\epsilon \rightarrow 0$ .

*Proof.*

$$\begin{aligned} & \int_Q (a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon), Du_\epsilon) dx \\ & = - \int_Q (\operatorname{div}(a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon)), u_\epsilon) dx = - \int_Q f u_\epsilon dx \\ & \rightarrow - \int_Q f u dx = \int_Q (a^*(u, Du), Du) dx. \quad \square \end{aligned}$$

Next for the second term on the r.h.s. of (17), using (6), we have

$$\begin{aligned} (19) \quad & C \left| \int_Q (a(T(x/\epsilon)\omega, u_\epsilon, \mathcal{P}_\epsilon) - a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon), \mathcal{P}_\epsilon - Du_\epsilon) dx \right| \\ & \leq \frac{C}{\delta_1} \left| \int_Q a(T(x/\epsilon)\omega, u_\epsilon, \mathcal{P}_\epsilon) - a(T(x/\epsilon)\omega, M_h u, \mathcal{P}_\epsilon) \right|^q dx + C\delta_1 \int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^p dx \\ & \leq \frac{C}{\delta_1} \sum_i \int_{Q_i} \nu(|u_\epsilon - \eta_i|)^q (1 + |\xi_i|^p) dx \\ & \quad + \frac{C}{\delta_1} \sum_i \int_{Q_i} \nu(|u_\epsilon - \eta_i|)^q (1 + |w_{\eta_i, \xi_i}|^p) dx + C\delta_1 \int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^p dx, \end{aligned}$$

where  $\nu(r)$  is a continuity modulus defined earlier (see (6)). Here we have used the uniform boundedness of  $Du_\epsilon$  as well as  $u_\epsilon$  in  $L^p(Q)^d$  and  $L^p(Q)$ , respectively. The first term on the r.h.s. converges to  $\int_Q \nu(|u - M_h u|)^q (1 + |M_h Du|^p) dx$  by Lemma 3.4. For the second term, using Meyers' theorem (cf. Lemma 3.6), we obtain that  $w_{\eta_i, \xi_i}$  is bounded in  $L^{p+\alpha}(Q_i)^d$ ,  $\alpha > 0$ . Thus, using Lemma 3.4, we have that the second term for each  $i$  converges to  $\int_{Q_i} \nu(|u - \eta_i|)^q (1 + |w_{\eta_i, \xi_i}|^p) dx$ , which is not

greater than  $\int_{Q_i} \nu(|u - \eta_i|)^q (1 + |\eta_i|^p + |\xi_i|^p) dx$ . Summing this over all  $i$ , we get  $\int_Q \nu(|u - M_h u|)^q (1 + |M_h u|^p + |M_h Du|^p) dx$ . Thus (19) is not greater than

$$\int_Q \nu(|u - M_h u|)^q (1 + |M_h u|^p + |M_h Du|^p) dx + C\delta_1 \int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^p dx.$$

With an appropriate choice of  $\delta_1$ , combining all the estimates, we have for (17) (cf. (16))

$$\begin{aligned} (20) \quad & \lim_{\epsilon \rightarrow 0} \int_Q |\mathcal{P}_\epsilon - Du_\epsilon|^p dx \\ & \leq C \left( \int_Q (a^*(M_h u, M_h Du), M_h Du) dx - \int_Q (a^*(M_h u, M_h Du), Du) dx \right. \\ & \quad \left. - \int_Q (a^*(u, Du), M_h Du) dx + \int_Q (a^*(u, Du), Du) dx \right) \\ & \quad + \int_Q \nu(|u - M_h u|)^q (1 + |M_h u|^p + |M_h Du|^p) dx. \end{aligned}$$

Next it is not difficult to show that the r.h.s. of (20) approaches zero as  $h \rightarrow 0$ . For this reason we write

$$\begin{aligned} (21) \quad & \int_Q (a^*(M_h u, M_h Du), M_h Du) dx - \int_Q (a^*(M_h u, M_h Du), Du) dx \\ & - \int_Q (a^*(u, Du), M_h Du) dx + \int_Q (a^*(u, Du), Du) dx \\ & = \int_Q (a^*(u, Du) - a^*(M_h u, M_h Du), Du - M_h Du) dx. \end{aligned}$$

Next, using the estimate  $|a^*(\eta_1, \xi_1) - a^*(\eta_2, \xi_2)| \leq C(1 + |\eta_1|^{p-1} + |\eta_2|^{p-1} + |\xi_1|^{p-1} + |\xi_2|^{p-1})\nu(|\eta_1 - \eta_2|) + C(1 + |\eta_1|^{p-1-\tilde{s}} + |\eta_2|^{p-1-\tilde{s}} + |\xi_1|^{p-1-\tilde{s}} + |\xi_2|^{p-1-\tilde{s}})|\xi_1 - \xi_2|^{\tilde{s}}$ ,  $0 < \tilde{r} \leq 1$  (see [17]), we can obtain that the r.h.s. of (21) converges to zero as  $h \rightarrow 0$ . Indeed,

$$\begin{aligned} (22) \quad & \int_Q (a^*(u, Du) - a^*(M_h u, M_h Du), Du - M_h Du) dx \\ & \leq C \int_Q (1 + |u|^{p-1} + |Du|^{p-1} + |M_h u|^{p-1} + |M_h Du|^{p-1}) \nu(|u - M_h u|) |Du - M_h Du| dx \\ & + C \int_Q (1 + |u|^{p-1-\tilde{s}} + |Du|^{p-1-\tilde{s}} + |M_h u|^{p-1-\tilde{s}} + |M_h Du|^{p-1-\tilde{s}}) |Du - M_h Du|^{\tilde{s}} dx. \end{aligned}$$

Using the Holder inequality, it can be easily shown that the second term here converges to zero as  $h \rightarrow 0$ . Since  $M_h u$  converges to  $u$  in  $L^p(Q)$  and  $M_h Du$  converges to  $Du$  in  $L^p(Q)^d$  from Lemma 3.4, the first term in (22) also converges to zero. Similarly one can show that the last term on the r.h.s. of (20) converges to zero as  $h \rightarrow 0$ . This completes the proof of Theorem 3.1.

As an example we consider the correctors for

$$(23) \quad \operatorname{div}(a(T(x/\epsilon)\omega)k_r(u_\epsilon)Du_\epsilon) = f.$$

We assume that the operator satisfies the conditions stated previously. In this case  $P(T(y)\omega, \eta, \xi) = \xi + w_{\eta, \xi}(T(y)\omega)$ , where  $w_{\eta, \xi} \in L^p_{pot}(\Omega)$  satisfies

$$-\operatorname{div}(a(T(x/\epsilon)\omega)k_r(\eta)(\xi + w_{\eta, \xi})) = 0.$$

Introducing a notation  $N$  such that  $w_{\eta, \xi}^i(\omega) = N_{ij}(\omega)\xi_j$ , we have the classical equation (see [13]) for  $N(\omega)$ , i.e.,  $a(\omega)(I + N) \in L^q_{sol}(\Omega)$ . Consequently, the correctors for (23) have the form

$$P(T(y)\omega, \eta, \xi) = \xi(I + N(T(y)\omega)).$$

From this we conclude that  $u$  satisfies

$$\operatorname{div}(a^*k_r(u)Du) = f,$$

where  $a^*$  is the homogenized tensor corresponding to a linear elliptic operator. The approximation for the gradient of the solution is defined by

$$P(T(x/\epsilon)\omega, M_h u, M_h Du) = M_h Du(I + N(T(x/\epsilon)\omega)).$$

Theorem 3.1 shows a way to compute an approximation for the gradient of  $u_\epsilon$ , although this computation is difficult since it involves the solution of the auxiliary problem. In the next section we will present the numerical computation of the oscillatory solution.

#### 4. Numerical computation of the homogenized solution.

**4.1. Numerical homogenization method.** Consider  $u_\epsilon \in W_0^{1,p}(Q)$ ,

$$(24) \quad -\operatorname{div}(a_\epsilon(x, u_\epsilon, Du_\epsilon)) + a_{0,\epsilon}(x, u_\epsilon, Du_\epsilon) = f,$$

where  $a_\epsilon(x, \eta, \xi)$  and  $a_{0,\epsilon}(x, \eta, \xi)$ ,  $\eta \in R$ ,  $\xi \in R^d$ , satisfy (4)–(6) and (7). As we mentioned in the introduction, the numerical homogenization procedure and its analysis can be studied for more general heterogeneities using  $G$ -convergence theory. The main idea of the numerical homogenization procedure is to find the homogenized solution without using the auxiliary problem. Consider a finite dimensional space over the standard triangular partitions  $K$  of  $Q = \bigcup K$ , and let

$$(25) \quad S^h = \{v_h \in C^0(\bar{Q}) : \text{the restriction } v_h \text{ is linear for each element } K \text{ and } v_h = 0 \text{ on } \partial Q\},$$

$\operatorname{diam}(K) \leq Ch$ . Here we assume that  $h \gg \epsilon$  is chosen for the approximation of the homogenized solution. The numerical homogenization procedure consists of finding an approximation,  $u_h \in S^h$ , of a homogenized solution  $u$  such that

$$(26) \quad (A_{\epsilon,h}u_h, v_h) = \int_Q f v_h dx,$$

where

$$(27) \quad (A_{\epsilon,h}u_h, v_h) = \sum_K \int_K ((a_\epsilon(x, \eta^{u_h}, Du_{\epsilon,h}), Dv_h) + a_{0,\epsilon}(x, \eta^{u_h}, Du_{\epsilon,h})v_h) dx.$$

Here  $u_{\epsilon,h}$  satisfies

$$(28) \quad -\operatorname{div}(a_\epsilon(x, \eta^{u_h}, Du_{\epsilon,h})) = 0 \text{ in } K,$$

$u_{\epsilon,h} = u_h$  on  $\partial K$ , and  $\eta^{v_h} (= M_h v_h) = \frac{1}{|K|} \int_K v_h dx$  in each  $K$ . Our numerical homogenization procedure consists of (26), (27), and (28). In some sense, (27) attempts to approximate  $\int_Q [(a^*(x, u_h, Du_h), Dv_h) + a_0^*(x, u_h, Du_h)v_h] dx$ , which is a finite element formulation of the homogenized equation. Note that solutions,  $u_h$ , of (26) depend on  $\epsilon$ , which we do not explicitly write because  $u_h \in S^h$ . We would like to point out that different boundary conditions can be chosen; e.g., one can use an oversampling technique [10], where the solution of the larger problem is used in the calculation of the solution of local problems. We have implemented and shown the advantages of an oversampling technique in our recent work [7]. In the next subsection we will show that the numerical homogenization approach can be considered as a generalization of MsFEM.

Next we briefly describe the numerical implementation of MsFEM for nonlinear elliptic problems. For each  $u_h = \sum_i \theta_i \phi_0^i(x) \in S^h$ , where  $\phi_0^i(x)$  is a basis in  $S^h$ , (26) is equivalent to solving

$$(29) \quad F(\theta) = b,$$

where  $F(\theta)$  is defined by (27) with  $v_h = \phi_0^i(x)$  and  $b_i = \int_Q f \phi_0^i(x) dx$ . Equation (29) can be solved using Newton's method or its modifications. This involves the inversion of the Jacobian corresponding to  $F(\theta)$ . When using MsFEM, the Jacobian is a matrix assembled on the coarse grid, which gives us the advantage in the computations.

The following convergence result will be shown.

**THEOREM 4.1.** *Let  $u_h$  and  $u$  be solutions of (26) and (2), respectively. Then*

$$(30) \quad \lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|u_h - u\|_{W_0^{1,p}(Q)} = 0$$

(up to a subsequence) under some nonrestrictive assumptions on  $a^*(x, \eta, \xi)$ .

*Remark 4.1.* Since the proof uses  $G$ -convergence theory, the limiting  $a^*$  (as well as  $a_0^*$ ) is not unique, and the convergence of the numerical solutions is up to a subsequence in  $\epsilon$ ; i.e.,  $u_h$  converges to a solution of a homogenized equation. We note that for the random homogeneous case the limiting operator is unique and the whole sequence converges. In later analysis, all the limits are taken up to a subsequence.

Note that because of the lack of scale separation, the above result cannot be improved, because there are all the scales  $\alpha(\epsilon)$ , such that  $\alpha(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$  are present. We have observed significant improvement in the numerical calculation when an oversampling technique is used for problems without scale separation. To show that  $u_{\epsilon,h}$  approximates  $u_\epsilon$  in  $W_0^{1,p}(Q)$  we will use the corrector results presented in the previous section.

**4.2. Numerical homogenization method and MsFEM.** To present the relation between the numerical homogenization approach and MsFEM we introduce the multiscale mapping,  $E^{MsFEM} : S^h \rightarrow V_\epsilon^h$ , a one-to-one operator which is constructed in the following way. For each  $v_h \in S^h$ ,  $v_{\epsilon,h}$  is the solution of

$$(31) \quad -\operatorname{div}(a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h})) = 0 \text{ in } K;$$

in each  $K$ ,  $v_{\epsilon,h} = v_h$  on  $\partial K$ , and  $\eta^{v_h} = \frac{1}{|K|} \int_K v_h dx$ . In [11] the authors introduce MsFEM, where a basis for  $V_\epsilon^h$  is constructed by mapping a basis of  $S^h$ . The extension of this approach to nonlinear problems yields a nonlinear space for the approximation of heterogeneities. Note that  $v_{\epsilon,h}$  are uniquely determined because (31) enjoys the monotonicity property. Now the numerical homogenization procedure can be written in the following way. Find  $u_h \in S^h$  (consequently,  $u_{\epsilon,h} = E^{MsFEM} u_h \in V_\epsilon^h$ ) such that

$$(32) \quad (A_{\epsilon,h} u_h, v_h) = \int_Q f v_h dx \quad \forall v_h \in S^h,$$

where  $A_{\epsilon,h}$  is given by (26). Later on we will show that  $Du_{\epsilon,h}$  approximates  $Du_\epsilon$  in  $L^p(Q)^d$ , assuming that the fluxes  $a_\epsilon(x, \eta, \xi)$  and  $a_{0,\epsilon}(x, \eta, \xi)$  are random homogeneous fields. Clearly, for periodic problems, (31) can be solved in a period of size  $\epsilon$  and extended periodically to  $K$ . This solution will approximate the solution of (31) and can be used in the construction of  $A_{\epsilon,h}$  and in setting up (32) (cf. HMM [6]). The convergence analysis for this case can be easily carried out using periodic correctors, and this will be presented elsewhere. Finally, we would like to note that one can adopt the oversampling technique [10] for nonlinear multiscale finite element methods.

**4.3. Proof of Theorem 4.1.** The proof of Theorem 4.1 will be carried out in the following way. First we show the coercivity of  $A_{\epsilon,h}$  defined by (27). Next we study the limit as  $\epsilon \rightarrow 0$  of (26) and show that the solution of the limiting equation approximates homogenized solutions. For the sake of simplification of the proof, we assume  $\beta = p$  in (4).

LEMMA 4.2. *Let  $A_{\epsilon,h}$  be defined by (27). Then for sufficiently small  $h$ , there exists a constant  $C > 0$  such that for any  $v_h \in S^h$*

$$(A_{\epsilon,h} v_h, v_h) \geq C \|Dv_h\|_{p,Q}^p - C_1.$$

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

$$(33) \quad -\operatorname{div} a_\epsilon(x, \eta^{v_h}, D\tilde{v}_{\epsilon,h} + Dv_h) = 0 \quad \text{in } K.$$

Using (33) and applying Green's theorem and (7), we have the following estimate:

$$\begin{aligned} (A_{\epsilon,h} v_h, v_h) &= \sum_K \int_K [(a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h}), Dv_h) + a_{0,\epsilon}(x, \eta^{v_h}, Dv_{\epsilon,h}) v_h] dx \\ &= \sum_K \int_K [(a_\epsilon(x, \eta^{v_h}, Dv_h + D\tilde{v}_{\epsilon,h}), Dv_h + D\tilde{v}_{\epsilon,h}) + a_{0,\epsilon}(x, \eta^{v_h}, Dv_{\epsilon,h}) v_h] dx \\ &= \sum_K \int_K [(a_\epsilon(x, \eta^{v_h}, Dv_h + D\tilde{v}_{\epsilon,h}), Dv_h + D\tilde{v}_{\epsilon,h}) + a_{0,\epsilon}(x, \eta^{v_h}, Dv_{\epsilon,h}) \eta^{v_h}] dx \\ &\quad + \sum_K \int_K a_{0,\epsilon}(x, \eta^{v_h}, Dv_{\epsilon,h}) (v_h - \eta^{v_h}) dx \\ &\geq C \sum_K \int_K |Dv_h + D\tilde{v}_{\epsilon,h}|^p dx - Ch \left( 1 + \sum_K \int_K |Dv_h|^p dx \right) - C_1. \end{aligned}$$

Here we have also used the fact that  $\int_K |\eta^{v_h}|^p dx \leq C \int_K |v_h|^p dx$ . Next we will show that

$$\sum_K \int_K |Dv_h + D\tilde{v}_{\epsilon,h}|^p dx = \sum_K \int_K |Dv_{\epsilon,h}|^p dx \geq C \sum_K \int_K |Dv_h|^p dx.$$

We note that  $v_h$  is piecewise linear on  $\partial K$  for triangular mesh, i.e.,  $v_{\epsilon,h}|_{\partial K} = v_h = \beta + (Dv_h, x - x_0)$ , for some constants  $\beta$  and  $x_0$  independent of  $Dv_h$ . We set  $\tilde{v}_{\epsilon,h} = v_{\epsilon,h} - \beta$ . Then, by change of variable and homogeneity argument and applying the trace theorem, we have

$$\begin{aligned} \sum_K \int_K |Dv_{\epsilon,h}|^p dx &\geq C \sum_K \frac{h^d}{h^p} \int_{K_r} |D_y \tilde{v}_{\epsilon,h}|^p dy \\ &\geq C \sum_K \frac{h^d}{h^p} \int_{\partial K_r} |(Dv_h, y h)|^p dy = C \sum_K h^d |Dv_h|^p C(e_{Dv_h}), \end{aligned}$$

where  $K_r$  is a reference triangle of size of order 1,  $e_{Dv_h}$  is the unit vector in the direction of  $Dv_h$ , and

$$C(e_{Dv_h}) = \int_{\partial K_r} |(e_{Dv_h}, y)|^p dy.$$

Here we have used the trace inequality,  $\|u\|_{L^p(\partial Q)} \leq C \|u\|_{W^{1,p}(Q)}$ , and taken into account the equivalence of finite dimensional norms for every  $h$ . One can further show that  $C(e_\xi)$  is bounded from below independent of  $\xi$  and  $h$ . By contradiction, suppose that the claim is not true. Then there exists a sequence  $\{e_{\xi_n}\}$  which has a subsequence (denoted by the same notation) such that  $e_{\xi_n} \rightarrow e_*$  and  $C(e_{\xi_n}) \rightarrow 0$  as  $n \rightarrow \infty$ . Since  $C(e_\xi)$  is continuous, it follows that  $C(e_*) = 0$ . This further implies that  $(e_*, y) = 0$  on  $\partial K_r$ , and hence  $e_* = 0$ . This is a contradiction.  $\square$

Next we show that  $A_{\epsilon,h}$  is equicontinuous for any  $h$  in any compact set.

LEMMA 4.3. *For any  $v_h \in S^h$  and  $w_h \in S^h$  in a compact set we have*

$$\|A_{\epsilon,h} v_h - A_{\epsilon,h} w_h\|^p \leq C \left( \sum_K \int_K (|D(v_h - w_h)|^p + \nu(|\eta^{v_h} - \eta^{w_h}|)) dx \right)^{1/p},$$

where  $C$  does not depend on  $\epsilon$ .

Since this result is for fixed  $h$  (i.e., finite dimensional), we do not specify the norm.

*Proof.*

$$\begin{aligned} (34) \quad \|A_{\epsilon,h} v_h - A_{\epsilon,h} w_h\| &= \sum_K \int_K |a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h}) - a_\epsilon(x, \eta^{w_h}, Dw_{\epsilon,h})| dx \\ &\quad + \sum_K \int_K |a_{0,\epsilon}(x, \eta^{v_h}, Dv_{\epsilon,h}) - a_{0,\epsilon}(x, \eta^{w_h}, Dw_{\epsilon,h})| dx. \end{aligned}$$

Next we will estimate the first term on the r.h.s. of (34). The estimate for the second

term is analogous.

$$\begin{aligned}
 & \sum_K \int_K |a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h}) - a_\epsilon(x, \eta^{w_h}, Dw_{\epsilon,h})| dx \\
 & \leq C \sum_K \int_K (1 + |\eta^{v_h}| + |\eta^{w_h}| + |Dv_{\epsilon,h}| + |Dw_{\epsilon,h}|)^{p-1} \nu(|\eta^{v_h} - \eta^{w_h}|) \\
 (35) \quad & + C \sum_K \int_K (1 + |\eta^{v_h}| + |\eta^{w_h}| + |Dv_{\epsilon,h}| + |Dw_{\epsilon,h}|)^{p-1-s} |Dv_{\epsilon,h} - Dw_{\epsilon,h}|^s \\
 & \leq C \left( \sum_K \int_K \nu(|\eta^{v_h} - \eta^{w_h}|)^p dx \right)^{1/p} + C \left( \sum_K \int_K |Dv_{\epsilon,h} - Dw_{\epsilon,h}|^p \right)^{1/p}.
 \end{aligned}$$

Here we have used the Cauchy inequality along with the facts that  $\|Dv_{\epsilon,h}\|_{p,K} \leq C\|Dv_h\|_{p,K}$ ,  $\|Dw_{\epsilon,h}\|_{p,K} \leq C\|Dw_h\|_{p,K}$ , and  $\|Dv_h\|_{p,Q} \leq C$ ,  $\|Dw_h\|_{p,Q} \leq C$ . It remains to estimate the second term on the r.h.s. of (35).

$$\begin{aligned}
 (36) \quad & \sum_K \int_K |Dv_{\epsilon,h} - Dw_{\epsilon,h}|^p \\
 & \leq C \sum_K \int_K (a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h}) - a_\epsilon(x, \eta^{v_h}, Dw_{\epsilon,h}), Dv_{\epsilon,h} - Dw_{\epsilon,h}) dx \\
 & \leq \sum_K \int_K (a_\epsilon(x, \eta^{v_h}, Dw_{\epsilon,h}) - a_\epsilon(x, \eta^{w_h}, Dw_{\epsilon,h}), Dv_{\epsilon,h} - Dw_{\epsilon,h}) dx \\
 & \quad + \sum_K \int_K (a_\epsilon(x, \eta^{w_h}, Dw_{\epsilon,h}) - a_\epsilon(x, \eta^{v_h}, Dw_{\epsilon,h}), Dv_{\epsilon,h} - Dw_{\epsilon,h}) dx \\
 & \leq C \sum_K \int_K (a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h}) - a_\epsilon(x, \eta^{w_h}, Dw_{\epsilon,h}), Dv_h + D\tilde{v}_{\epsilon,h} - Dw_h - D\tilde{w}_{\epsilon,h}) dx \\
 & \quad + C \sum_K \int_K (a_\epsilon(x, \eta^{w_h}, Dw_{\epsilon,h}) - a_\epsilon(x, \eta^{v_h}, Dw_{\epsilon,h}), Dv_{\epsilon,h} - Dw_{\epsilon,h}) dx \\
 & \leq C \sum_K \int_K (a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h}) - a_\epsilon(x, \eta^{w_h}, Dw_{\epsilon,h}), Dv_h - Dw_h) dx \\
 & \quad + C \sum_K \int_K \nu(|\eta^{w_h} - \eta^{v_h}|)^p dx \\
 & \leq C \left( \sum_K \int_K |Dv_h - Dw_h|^p dx \right)^{1/p} + C \sum_K \int_K \nu(|\eta^{w_h} - \eta^{v_h}|)^p dx.
 \end{aligned}$$

Here we have used Holder and Cauchy inequalities along with the facts that  $\|Dv_{\epsilon,h}\|_{p,K} \leq C\|Dv_h\|_{p,K}$ ,  $\|Dw_{\epsilon,h}\|_{p,K} \leq C\|Dw_h\|_{p,K}$ , and  $\|Dv_h\|_{p,Q} \leq C$ ,  $\|Dw_h\|_{p,Q} \leq C$ , and that  $v_{\epsilon,h} = v_h + \tilde{v}_{\epsilon,h}$ , where  $\tilde{v}_{\epsilon,h} \in W_0^{1,p}(K)$  satisfies  $-\operatorname{div} a_\epsilon(x, \eta^{v_h}, Dv_h + D\tilde{v}_{\epsilon,h}) = 0$ . The estimates (35) and (36) give us the estimate for the first term of the r.h.s. of (34). A similar estimate for the second term can be obtained in a very analogous manner.  $\square$

The coercivity and continuity of  $A_{\epsilon,h}$  guarantee the existence of a solution for the

discrete equation

$$(37) \quad (A_{\epsilon,h}u_{\epsilon,h}, w_h) = \int_Q f w_h dx.$$

LEMMA 4.4. *For any  $v_h \in S^h$  and  $w_h \in S^h$*

$$\lim_{\epsilon \rightarrow 0} (A_{\epsilon,h}v_h, w_h) = (A_h v_h, w_h)$$

(up to a sub-sequence), where the r.h.s. is defined as

$$(A_h v_h, w_h) = \sum_K \int_K [(a^*(x, \eta^{v_h}, Dv_0), Dw_h) + a_0^*(x, \eta^{v_h}, Dv_0)w_h] dx$$

and  $v_0$  is the solution of  $v_0 - v_h \in W_0^{1,p}(K)$ ,

$$-\operatorname{div}(a^*(x, \eta^{v_h}, Dv_0)) = 0.$$

Here  $a^*(x, \eta, \xi)$  is a  $G$ -limit of the corresponding limit operator.

*Proof.* Using the theorem on  $G$ -convergence of arbitrary solutions [17, p. 87], we obtain that solutions  $v_{\epsilon,h}$  of (31) weakly converge to  $v_0$  in  $W^{1,p}(K)$ , and  $a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h})$  weakly converges to  $a^*(x, \eta^{v_h}, Dv_0)$  in  $L^q(K)^d$ , and  $a_{0,\epsilon}(x, \eta^{v_h}, Dv_{\epsilon,h})$  weakly converges to  $a_0^*(x, \eta^{v_h}, Dv_0)$  in  $L^q(K)$  (up to a subsequence), where  $a^*(x, \eta, \xi)$  and  $a_0^*(x, \eta, \xi)$  are the fluxes corresponding to a  $G$ -limit of the original operators. Thus,

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} (A_{\epsilon,h}v_h, w_h) &= \lim_{\epsilon \rightarrow 0} \sum_K \int_K [(a_\epsilon(x, \eta^{v_h}, Dv_{\epsilon,h}), Dw_h) + a_{0,\epsilon}(x, \eta^{v_h}, Dv_{\epsilon,h})w_h] dx \\ &= \sum_K \int_K [(a^*(x, \eta^{v_h}, Dv_0), Dw_h) + a_0^*(x, \eta^{v_h}, Dv_0)w_h] dx = (A_h v_h, w_h). \quad \square \end{aligned}$$

It can be easily shown that  $A_h$  is coercive for small  $h$ . Since  $A_{\epsilon,h}$  is equicontinuous in any compact set, the results of Lemma 4.4 hold for any  $v_h \in S^h$  and  $w_h \in S^h$  that are uniformly bounded (finite dimensional). Thus, taking the limit  $\epsilon \rightarrow 0$  of (37) (up to a subsequence), we obtain

$$(A_h u_h, w_h) = \int_Q f w_h dx.$$

From Lemma 4.2 and the continuity of  $A_h$  (which can be easily verified) it follows that  $u_h$  exists and is uniformly bounded in  $W_0^{1,p}(Q)$ , and thus  $u_h \rightarrow u$  (up to a subsequence) weakly in  $W_0^{1,p}(Q)$ . Our task is to show that  $u$  is a solution of the homogenized equation. The following lemma is needed for this purpose.

LEMMA 4.5. *Assume that  $v_h \in S^h$  and  $Dv_h$  are uniformly bounded in  $L^{p+\alpha}(Q)^d$  (with  $\alpha > 0$ ) and  $w_h \in S^h$  and  $Dw_h$  are uniformly bounded in  $L^p(Q)^d$ . Then*

$$\lim_{h \rightarrow 0} (A_h v_h - A^* v_h, w_h) = 0,$$

where

$$(A^* v_h, w_h) = \sum_K \int_K [(a^*(x, v_h, Dv_h), Dw_h) + a_0^*(x, v_h, Dv_h)w_h] dx.$$



*Proof.*

$$(38) \quad (A_h v_h - A^* v_h, w_h) = \sum_K \int_K [(a^*(x, \eta^{v_h}, Dv_0) - a^*(x, v_h, Dv_h), Dw_h) + (a_0^*(x, \eta^{v_h}, Dv_0) - a_0^*(x, v_h, Dv_h))w_h] dx.$$

Next we will show that the first and second terms on the r.h.s. of (38) converge to zero. For the first term we have

$$(39) \quad \begin{aligned} & \sum_K \int_K (a^*(x, \eta^{v_h}, Dv_0) - a^*(x, v_h, Dv_h), Dw_h) dx \\ & \leq C \sum_K \int_K \nu(|v_h - \eta^{v_h}|)(1 + |\eta^{v_h}|^{p-1} + |v_h|^{p-1} + |Dv_0|^{p-1} + |Dv_h|^{p-1}) |Dw_h| dx \\ & \quad + C \sum_K \int_K (1 + |\eta^{v_h}| + |v_h| + |Dv_0| + |Dv_h|)^{p-1-s} |Dv_h - Dv_0|^s |Dw_h| dx \\ & \leq C \sum_K \left( \int_K \nu(|v_h - \eta^{v_h}|)^q (1 + |Dv_h|^p) dx \right)^{1/q} \left( \int_K |Dw_h|^p dx \right)^{1/p} \\ & \quad + C \sum_K \left( \int_K (1 + |Dv_h|^p) dx \right)^{(p-q)s/pq} \left( \int_K |D(v_h - v_0)|^p dx \right)^{s/p} \left( \int_K |Dw_h|^p dx \right)^{1/p} \\ & = C \left( \int_Q \nu(|v_h - \eta^{v_h}|)^q (1 + |Dv_h|^p) dx \right)^{1/q} + C \left( \int_Q |D(v_h - v_0)|^p dx \right)^{s/p}. \end{aligned}$$

Here we have used the Cauchy inequality along with the facts that  $\|Dw_h\|_{p,Q} \leq C$ ,  $\int_K |Dv_0|^p dx \leq C \int_K |Dv_h|^p dx$ ,  $\int_K |v_h|^p dx \leq C \int_K |Dv_h|^p dx$ , and  $\int_K |\eta^{v_h}|^p dx \leq C \int_K |Dv_h|^p dx$ . Next we will show that  $\|Dv_h - Dv_0\|_{p,Q} \rightarrow 0$  as  $h \rightarrow 0$  under some assumptions regarding the regularity of  $a^*(x, \eta, \xi)$  with respect to spatial variables. Moreover, this convergence is uniform for a uniformly bounded family of  $Dv_h$ . Define  $\overline{a^*}_K(x, \eta, \xi)$  as a piecewise constant function on each  $K$  and  $\eta, \xi$  defined in each  $K$  by

$$\overline{a^*}_K(\eta, \xi) = \frac{1}{|K|} \int_K a^*(x, \eta, \xi).$$

We assume that in each  $K$

$$(40) \quad |a^*(x, \eta, \xi) - \overline{a^*}_K(\eta, \xi)| \leq \alpha_h (1 + |\eta|^{p-1} + |\xi|^{p-1}),$$

where  $\alpha_h$  is a generic sequence such that  $\alpha_h \rightarrow 0$  as  $h \rightarrow 0$  and is independent of  $K$ . For example, this condition is satisfied if  $a^*(x, \eta, \xi)$  is a Holder function with respect to spatial variables. Note that for random homogeneous operators, (40) trivially holds because  $a^*$  is independent of  $x$ .

Then

$$\begin{aligned}
\|Dv_h - Dv_0\|_{p,Q}^p &\leq C \sum_K \int_K (a^*(x, \eta^{v_h}, Dv_h) - a^*(x, \eta^{v_h}, Dv_0), Dv_h - Dv_0) dx \\
&= C \sum_K \int_K (a^*(x, \eta^{v_h}, Dv_h) - \bar{a}^*_K(\eta^{v_h}, Dv_h), Dv_h - Dv_0) dx \\
&\quad + C \sum_K \int_K (\bar{a}^*_K(\eta^{v_h}, Dv_h) - a^*(x, \eta^{v_h}, Dv_0), Dv_h - Dv_0) dx \\
&= C \sum_K \int_K (a^*(x, \eta^{v_h}, Dv_h) - \bar{a}^*_K(\eta^{v_h}, Dv_h), Dv_h - Dv_0) dx \\
&\leq C\alpha_h \sum_K \int_K (1 + |\eta^{v_h}|^{p-1} + |Dv_h|)^{p-1} |Dv_h - Dv_0| dx \\
&\leq C\alpha_h (1 + \|Dv_h\|_{p,Q}^{p/q}) \|Dv_h - Dv_0\|_{p,Q}.
\end{aligned}$$

From here it follows that  $\|Dv_h - Dv_0\|_{p,Q} \rightarrow 0$  as  $h \rightarrow 0$ .

We note that the r.h.s. of (39) converges to zero because  $\|Dv_h - Dv_0\|_{p,Q} \rightarrow 0$  as  $h \rightarrow 0$  and because of Lemma 3.4. Thus, the first term on the r.h.s. of (38) converges to zero.

For the second term on the r.h.s. of (38) we have

$$\begin{aligned}
(41) \quad &\sum_K \int_K (a_0^*(x, \eta^{v_h}, Dv_0) - a_0^*(x, v_h, Dv_h)) w_h dx \\
&\leq C \sum_K \int_K \nu(|v_h - \eta^{v_h}|) (1 + |\eta^{v_h}|^{p-1} + |v_h|^{p-1} + |Dv_0|^{p-1} + |Dv_h|^{p-1}) |w_h| dx \\
&\quad + C \sum_K \int_K (1 + |\eta^{v_h}| + |v_h| + |Dv_0| + |Dv_h|)^{p-1-s} |Dv_h - Dv_0|^s |w_h| dx.
\end{aligned}$$

Clearly one can do the same manipulations as the those for the first term of the r.h.s. of (38) to show that the r.h.s. of (41) converges to zero as  $h \rightarrow 0$ .  $\square$

For the proof of Theorem 4.1 we assume that  $Du_h$  is uniformly bounded in  $L^{p+\alpha}(Q)^d$  for some  $\alpha > 0$ . One can assume additional nonrestrictive regularity assumptions [16] for input data and obtain Meyers-type estimates,  $\|Du\|_{p+\alpha,Q} \leq C$ , for the homogenized solutions. In this case it is reasonable also to assume that the discrete solutions are uniformly bounded in  $L^{p+\alpha}(Q)^d$ . We have obtained results on Meyers-type estimates for our approximate solutions in the case  $p = 2$  (see [8]). We are currently studying the generalization of these results to arbitrary  $p$ . One can impose different kinds of assumptions for which the Lemma 4.5 holds without assuming that  $Du_h$  is uniformly bounded in  $L^{p+\alpha}(Q)^d$ , e.g.,

$$|a^*(x, \eta, \xi) - a^*(s, \eta', \xi)| \leq C(1 + |\eta|^{p-1} + |\eta'|^{p-1} + |\xi|^{p-1-r}) |\eta - \eta'|^r$$

( $0 < r < 1$ ).

To conclude the proof of Theorem 4.1 we note that  $u_h \rightarrow u$  (up to a sub-sequence) weakly in  $W_0^{1,p}(Q)$ , and our goal is to show that  $u$  is a solution of the homogenized equation. Using Lemma 4.5, we obtain that

$$(A_h u_h - A^* u_h, v_h) = \int_Q f v_h dx - (A^* u_h, v_h).$$

Thus it follows from Lemma 4.5 that  $A^*u_h \rightarrow f$  weakly in  $W^{-1,p}(Q)$ . Moreover, using Lemma 4.5, we obtain that  $(A^*u_h, u_h) \rightarrow \int_Q f v_h dx$ . Since the operator  $A^*$  is of type  $M$  (see [20]), we obtain that  $A^*u = f$ ; i.e.,  $u$  is a solution of a homogenized equation. Moreover,  $A^*$  is also of type  $S_+$  (see [21]), which allows us to state that  $u_h \rightarrow u$  strongly in  $W_0^{1,p}(Q)$ .

*Remark 4.2.* We would like to note that in the periodic and random homogeneous cases Theorem 4.1 holds in the limit  $\epsilon/h \rightarrow 0$ ; i.e.,  $h = h(\epsilon) \gg \epsilon$ . This will be presented elsewhere.

*Remark 4.3.* Finally we would like to note that Theorem 4.1 is proved under the assumptions (40) and  $\|Du_h\|_{p+\alpha,Q} \leq C$ ,  $\alpha > 0$ . The latter has been shown for  $p = 2$  in [8].

**4.4. Approximation of the oscillations.** In order to approximate solutions  $u_\epsilon$  in the  $W^{1,p}$ -norm, we assume  $a_\epsilon(x, \eta, \xi) = a(T(x/\epsilon), \eta, \xi)$  and  $a_{0,\epsilon}(x, \eta, \xi) = a_0(T(x/\epsilon), \eta, \xi)$ . Then the following theorem holds.

THEOREM 4.6.

$$\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|D(u_{\epsilon,h} - u_\epsilon)\|_{p,Q} = 0,$$

where  $u_{\epsilon,h} = E^{MsFEM} u_h$ , defined by (31) (or (28) in each  $K$ ).

*Proof.* Because of Theorem 3.1 we need to show only that

$$\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|Du_{\epsilon,h} - P(T(x/\epsilon)\omega, M_h u, M_h Du)\|_{p,Q} = 0.$$

Similarly,

$$(42) \quad \lim_{\epsilon \rightarrow 0} \|Du_{\epsilon,h} - P(T(x/\epsilon)\omega, M_h u_h, M_h Du_h)\|_{p,K} = 0.$$

Equation (42) follows from the fact that  $-\operatorname{div}(a^*(\eta^{u_h}, D_x u_h)) = 0$ , i.e., the homogenized solution for  $u_{\epsilon,h}$  is  $u_h$ . Consequently,

$$\lim_{\epsilon \rightarrow 0} \|Du_{\epsilon,h} - P(T(x/\epsilon)\omega, M_h u_h, M_h Du_h)\|_{p,Q} = 0.$$

It remains to show that

$$\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|P(T(x/\epsilon)\omega, M_h u_h, M_h Du_h) - P(T(x/\epsilon)\omega, M_h u, M_h Du)\|_{p,Q} = 0.$$

To show this we need the estimate for  $\int_\Omega |P(\omega, \eta_1, \xi_1) - P(\omega, \eta_2, \xi_2)|^p d\mu(\omega)$ . Define  $P_1 = P(\omega, \eta_1, \xi_1)$  and  $P_2 = P(\omega, \eta_2, \xi_2)$ . Then

$$\begin{aligned} \int_\Omega |P_1 - P_2|^p d\mu(\omega) &\leq C \int_\Omega (a(\omega, \eta_1, P_1) - a(\omega, \eta_1, P_2), P_1 - P_2) d\mu(\omega) \\ &= \int_\Omega (a(\omega, \eta_1, P_1) - a(\omega, \eta_2, P_2), P_1 - P_2) d\mu(\omega) \\ &\quad + \int_\Omega (a(\omega, \eta_2, P_2) - a(\omega, \eta_1, P_2), P_1 - P_2) d\mu(\omega) \\ &\leq \int_\Omega (a(\omega, \eta_1, P_1) - a(\omega, \eta_2, P_2), \xi_1 - \xi_2) d\mu(\omega) \\ &\quad + \frac{C}{\delta_1} \int_\Omega (1 + |\eta_1|^p + |\eta_2|^p + |P_2|^p) \nu(|\eta_1 - \eta_2|) d\mu(\omega) \\ &\quad + C\delta_1 \int_\Omega |P_1 - P_2|^p d\mu(\omega). \end{aligned}$$

Choosing  $\delta_1$  appropriately small and using (10), we have

$$(43) \quad \int_{\Omega} |P_1 - P_2|^p d\mu(\omega) \leq (a^*(\eta_1, \xi_1) - a^*(\eta_2, \xi_2), \xi_1 - \xi_2) + C \int_{\Omega} (1 + |\eta_1|^p + |\eta_2|^p + |P_2|^p) \nu(|\eta_1 - \eta_2|) d\mu(\omega).$$

Using (43), we have

$$\begin{aligned} & \lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|P(T(x/\epsilon)\omega, M_h u_h, M_h Du_h) - P(T(x/\epsilon)\omega, M_h u, M_h Du)\|_{p,Q} \\ & \leq \lim_{h \rightarrow 0} \sum_K \int_K (a^*(M_h u_h, M_h Du_h) - a^*(M_h u, M_h Du), M_h Du_h - M_h Du) dx \\ & \quad + C \lim_{h \rightarrow 0} \sum_K \int_K (1 + |M_h u_h| + |M_h u| + |M_h Du|)^p \nu(|M_h u_h - M_h u|) dx. \end{aligned}$$

The r.h.s. of (4.4) converges to zero, which can be established in a manner similar to the convergence analysis of the r.h.s. of (20).  $\square$

**5. Numerical results.** Our first numerical example is a nonlinear convection diffusion equation in two dimensions:

$$(44) \quad \frac{1}{\epsilon} v(T(x/\epsilon)\omega) \cdot DF(u_\epsilon) - d\Delta u_\epsilon = f,$$

where  $\operatorname{div} v = 0$ . Assuming that there exists a homogeneous stream function  $\mathcal{H}(T(x/\epsilon)\omega)$ ,

$$\mathcal{H} = \begin{pmatrix} 0 & H(T(x/\epsilon)\omega) \\ -H(T(x/\epsilon)\omega) & 0 \end{pmatrix},$$

such that  $\operatorname{div} \mathcal{H} = v$ , we obtain

$$\operatorname{div}(-d\delta_{ij} Du_\epsilon + \mathcal{H}(T(x/\epsilon)\omega) DF(u_\epsilon)) = f$$

or

$$-\operatorname{div}(a(T(x/\epsilon)\omega, u_\epsilon) Du_\epsilon) = f,$$

where

$$a = \begin{pmatrix} -d & H(T(x/\epsilon)\omega) F'(u) \\ -H(T(x/\epsilon)\omega) F'(u) & -d \end{pmatrix}.$$

We assume that  $a$  satisfies the assumptions imposed in previous sections. The auxiliary problem is defined as follows:  $w_{\eta, \xi}(\omega) \in V_{pot}^p$  is the solution of

$$\operatorname{div}(a(T(y)\omega, \eta)(\xi + w_{\eta, \xi}(T(y)\omega))) = 0.$$

Introducing  $N_\eta^j(T(y)\omega) \in V_{pot}^p$  such that  $w_{\eta, \xi}^i(T(y)\omega) = N_\eta^{ij}(T(y)\omega)\xi_j$ , we have

$$\operatorname{div}(a(T(y)\omega, \eta)(I + N_\eta(T(y)\omega))) = 0,$$

where  $I$  is the identity matrix. Using  $w_{\eta,\xi}$ , we can compute the homogenized operator that is given by

$$\operatorname{div}(a^*(u)Du) = f,$$

where  $a_{ij}^*(\eta) = -d\delta_{ij} + \langle H_{ik}F'(\eta)N_{\eta}^{kj} \rangle$ . Here we have taken into account that  $\langle N \rangle = 0$  since  $N \in V_{pot}^p$ . The term  $\langle H_{ik}F'(\eta)N_{\eta}^{kj} \rangle$  can be regarded as an enhanced diffusion due to heterogeneous convection, similar to the linear case [9]. A numerical corrector is defined as

$$\mathcal{P}_{\epsilon} = M_h Du(I + w_{M_h u}(T(x/\epsilon)\omega)).$$

Next we present numerical examples where the enhanced diffusivity is approximately computed locally. It is more transparent for this purpose to use a parabolic equation,

$$(45) \quad \frac{\partial u_{\epsilon}}{\partial t} + \frac{1}{\epsilon} v(T(x/\epsilon)\omega) \cdot DF(u_{\epsilon}) = d\Delta u_{\epsilon}.$$

Using general  $G$ -convergence theory, we have the following equation for the homogenized solution:

$$(46) \quad \frac{\partial u}{\partial t} = \operatorname{div}(a^*(u)u),$$

where  $a^*(\eta)$  is the homogenized operator derived from the elliptic problem shown above. In particular,  $a_{ij}^* = d\delta^{ij} + a_{ij}^c$ , where  $a_{ij}^c(\eta) = -\langle H_{ik}F'(\eta)N_{\eta}^{kj} \rangle$  is the enhanced diffusion due to nonlinear heterogeneous convection. It can be shown that the corrector has the same form as in the elliptic case

$$P(T(x/\epsilon)\omega, M_h u(t, x), M_h Du(t, x)) = M_h Du(t, x)(I + w_{M_h u(t, x)}(T(x/\epsilon)\omega)),$$

i.e., all the time dependence is in the homogenized solution. The proof is the same as in the elliptic case.

To illustrate the significance of the enhanced diffusion, we present some numerical examples. Numerical tests are performed using the finite element method. First we present the total diffusivity as a function of  $\eta$  (i.e., average of the solution) for two different heterogeneous velocity fields given by the stream functions  $H = \sin(2\pi y/\epsilon) + \sin(2\sqrt{2}\pi y/\epsilon)$ . We take  $\epsilon = 0.1$  and  $d = 0.1$  (molecular diffusion). The flux function is chosen to be the Buckley–Leverett function  $F(u) = u^2/(u^2 + 0.2(1-u)^2)$  motivated by porous media flows. The enhanced diffusion is computed by solving the problem in a unit square, and thus it is only an approximate value of it. In Figure 1 we plot the total diffusivity. The left plot in this figure represents the total diffusivity in the horizontal direction (along the layers), and the right plot represents the total diffusivity in the vertical direction. Clearly, the diffusion is enhanced dramatically in the horizontal direction, that is, along the convection (note the ten-fold difference between the  $y$ -axis scales). As we see for  $\eta \approx 0.4$ , there is a 15-fold increase in the diffusion relative to molecular diffusion,  $d$ . Moreover, since  $F'(0) = F'(1) = 0$ , there is no enhancement if  $\eta = 0$  or  $\eta = 1$  (this corresponds to pure phases). For the cellular flow,  $H(x, y) = \sin(2\pi y/\epsilon)\sin(2\pi x/\epsilon)$ , we obtain isotropic diffusion, which is shown in Figure 2.

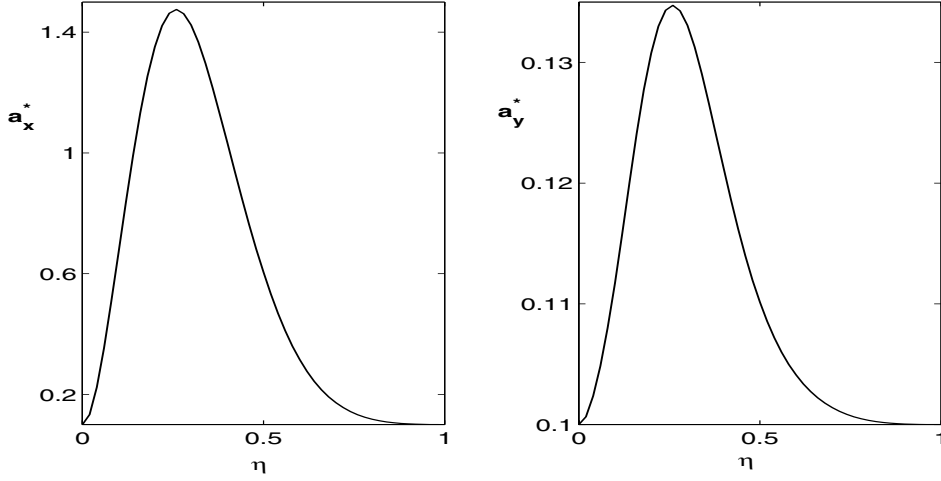


FIG. 1. Horizontal (left) and vertical (right) effective diffusivity for the layered media with stream function  $H(x, y) = \sin(2\pi y/\epsilon)$  and flux function  $F(u) = u^2/(u^2 + 0.2(1 - u)^2)$ .

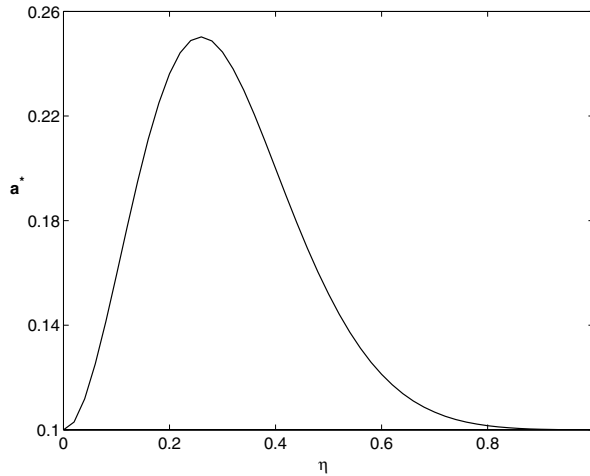


FIG. 2. Effective diffusivity for the isotropic media with stream function  $H(x, y) = \sin(2\pi y/\epsilon) \times \sin(2\pi x/\epsilon)$  and flux function  $F(u) = u^2/(u^2 + 0.2(1 - u)^2)$ .

The next set of numerical examples is designed to compare the solutions of the original (fine scale equation) with the solutions of the equations obtained using numerical homogenization with and without enhanced diffusion. Our goal here is to illustrate the importance of nonlinear enhanced diffusion. We consider

$$(47) \quad \frac{\partial u_\epsilon}{\partial t} + \frac{1}{\epsilon} v_\epsilon \cdot DF(u_\epsilon) = d\Delta u_\epsilon$$

in a unit square domain with boundary and initial conditions as follows:  $u_\epsilon = 1$  at the inlet ( $x_1 = 0$ ),  $u_\epsilon = 0$  at the outlet ( $x_1 = 1$ ), and no flow boundary conditions on the lateral sides  $x_2 = 0$  and  $x_2 = 1$ . We have tested various heterogeneous fields for the velocity, and we present here a result for the layered flow,  $H = \sin(2\pi y/\epsilon)$ .

In Figure 3 we plot the average (over the whole domain) of the solutions of (47) ( $\frac{1}{Q} \int_Q u_\epsilon(x, t) dx$ ) as a function of time. We compare the fine scale solution with the

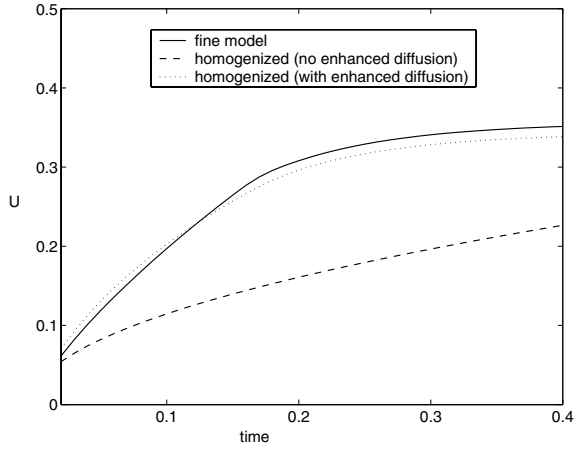


FIG. 3. Comparison of the average  $u$  over the whole domain for three problems: (1) fine scale (designated with a solid line), (2) homogenized solution with no enhanced diffusion (designated with a dashed line), and (3) homogenized solution with enhanced diffusion (designated with a dotted line). In this case  $H(x, y) = \sin(2\pi y/\epsilon)$  and  $F(u) = u^2/(u^2 + 0.2(1 - u)^2)$ .

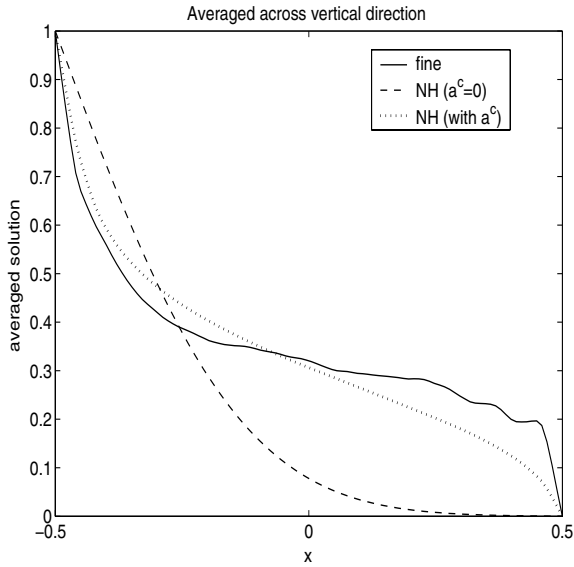


FIG. 4. Vertical average (across the heterogeneities) of the solution for the layered media with stream function  $H(x, y) = \sin(2\pi y/\epsilon)$  and flux function  $F(u) = u^2/(u^2 + 0.2(1 - u)^2)$  at time 0.4.

coarse scale solutions where the enhanced diffusivity is taken into account; i.e., it can be considered as a numerical homogenization procedure with one coarse block. We also consider the coarse scale solution where the enhanced diffusion is neglected, i.e.,  $u_t = d\Delta u$ . As we see from this figure, the solution computed with enhanced diffusion performs well and gives a reasonable approximation of the fine scale solution. On the other hand, the average solution that does not account for enhanced diffusion performs very poorly. In Figure 4 we plot the average along the horizontal direction ( $x_2$ )

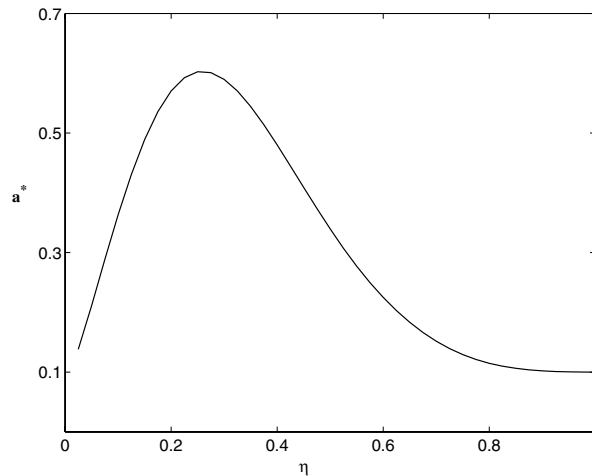


FIG. 5. *Effective diffusivity for the isotropic media with Gaussian stream function which has correlation lengths  $l_x = l_y = 0.1$ , mean zero, and variance 0.5. The flux function is chosen as  $F(u) = u^2/(u^2 + 0.2(1 - u)^2)$  and  $d = 0.1$  (see (44)).*

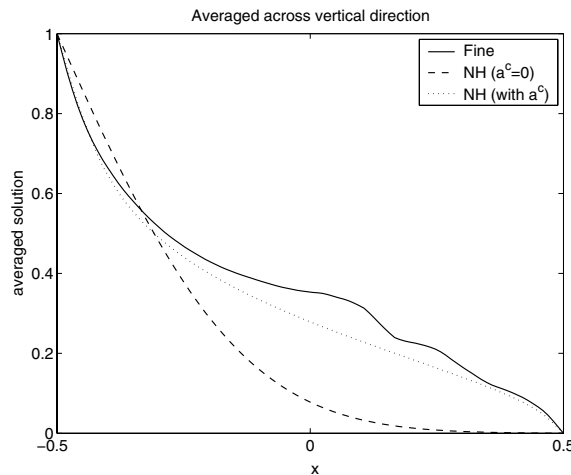


FIG. 6. *Average of the solution along the horizontal direction at  $t = 0.4$  for Gaussian stream function which has correlation lengths  $l_x = l_y = 0.1$ , mean zero, and variance 0.5. The flux function is chosen as  $F(u) = u^2/(u^2 + 0.2(1 - u)^2)$  and  $d = 0.1$  (see (44)).*

(across heterogeneities) of the solutions at time 0.4. The figure clearly indicates the importance of having enhanced diffusion in the homogenized setting of the problem. Next we present an example where the stream function  $H(x, y)$  is a realization of the random field with Gaussian distribution. To generate a realization of the random field with prescribed correlation lengths, we use GSLIB [5]. In particular,  $H(x, y)$  is a realization of a Gaussian field with correlation lengths  $l_x = l_y = 0.1$ , and with mean zero and variance 0.5. Here  $d = 0.1$  and  $F(u) = u^2/(u^2 + 0.2(1 - u)^2)$  are used in (44). In Figure 5 we plot the total diffusivity. As we see, the enhancement of the diffusion can be up to 6 times the molecular diffusion,  $d$ . Since the stream field is isotropic, it is sufficient to consider total diffusion in one direction. In Figure 6 we plot a cross section



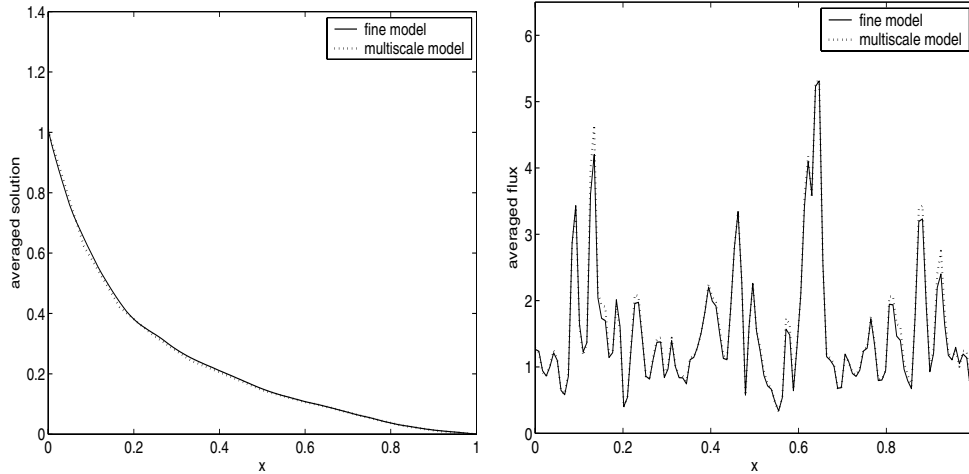


FIG. 7. Left: the solutions are averaged in the vertical direction. Right: the fluxes are averaged in the vertical direction.

of the solution at time 0.4. These results clearly indicate the importance of enhanced diffusion. For different realizations of the random field we have observed similar results. We note that this example can be easily generalized to nonlinear convection diffusion of more general form,  $\frac{1}{\epsilon}v(T(x/\epsilon)\omega) \cdot DF(u_\epsilon) - \text{div}(a(T(x/\epsilon)\omega, u_\epsilon, Du_\epsilon)) = f$ .

Finally, we consider an application of the numerical homogenization procedure to Richards equations,  $\text{div}(a_\epsilon(x, u_\epsilon)D_x u_\epsilon) = 0$ , 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 realization of a random field with the spherical variogram [5], correlation lengths  $l_x = 0.2$ ,  $l_y = 0.02$ , and variance  $\sigma = 1.5$ . Here  $\alpha_\epsilon(x)$  is chosen such that  $\alpha_\epsilon(x) = k_\epsilon(x) + \text{const}$  with the spatial average of 2. In Figure 7 we compare the solutions ( $u_\epsilon$ ) and the fluxes ( $-a_\epsilon(x, u_\epsilon)D_x u_\epsilon$ ) corresponding to this equation with boundary and initial conditions given as previously. The solutions are rescaled for comparison purposes. The solid line designates the fine scale model results computed on a  $120 \times 120$  grid, and the dotted line designates the coarse scale results computed using the numerical homogenization procedure on a  $12 \times 12$  coarse grid. These results demonstrate the robustness of our approach for anisotropic fields where  $h$  and  $\epsilon$  are nearly the same. For different realizations of the random field we have observed similar results. Currently, we are studying the application of the oversampling technique to the numerical homogenization procedure.

**Acknowledgments.** Efendiev would like to thank T. Hou for inspiring discussions. We would also like to acknowledge the anonymous reviewers for their helpful comments.

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. 329, AMS, Providence, RI, 2003, pp. 21–32.
- [2] I. BABUŠKA, G. CALOZ, AND J. E. OSBORN, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, SIAM J. Numer. Anal., 31 (1994), pp. 945–981.
- [3] F. BREZZI, *Interacting with the subgrid world*, in Numerical Analysis 1999 (Dundee), Chapman & Hall/CRC, Boca Raton, FL, 2000, pp. 69–82.

- [4] G. DAL MASO AND A. DEFRANCESCHI, *Correctors for the homogenization of monotone operators*, Differential Integral Equations, 3 (1990), pp. 1151–1166.
- [5] C. V. DEUTSCH AND A. G. JOURNEL, *GSLIB: Geostatistical Software Library and User's Guide*, 2nd ed., Oxford University Press, New York, 1998.
- [6] W. E AND B. ENGQUIST, *The heterogeneous multi-scale methods*, Comm. Math. Sci., 1 (2003).
- [7] Y. EFENDIEV, T. HOU, AND V. GINTING, *Multiscale finite element methods for nonlinear problems and their applications*, Comm. Math. Sci., submitted.
- [8] Y. EFENDIEV AND A. PANKOV, *Meyers type estimates for approximate solutions of nonlinear elliptic equations and their applications*, submitted to Numer. Math.; also available online at <http://www.math.tamu.edu/~yalchin.efendiev/ep-meyers-elliptic.ps>.
- [9] A. FANNJIANG AND G. PAPANICOLAOU, *Convection enhanced diffusion for periodic flows*, SIAM J. Appl. Math., 54 (1994), pp. 333–408.
- [10] T. Y. HOU AND X. H. WU, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
- [11] 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 (1999), pp. 913–943.
- [12] T. HUGHES, G. FEIJOO, L. MAZZEI, AND J. QUINCY, *The variational multiscale method—A paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg., 166 (1998), pp. 3–24.
- [13] V. V. JIKOV, S. M. KOZLOV, AND O. A. OLEINIK, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, New York, 1994.
- [14] 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, The Netherlands, 1976 (translated from the Russian by T. Ando).
- [15] 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), 2000, Vol. 33, pp. 43–59.
- [16] 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 (1975), pp. 121–136.
- [17] A. PANKOV, *G-convergence and Homogenization of Nonlinear Partial Differential Operators*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.
- [18] G. PAPANICOLAOU AND S. R. S. VARADHAN, *Boundary value problems with rapid oscillating random coefficients*, Seria Colloquia Mathematica Societatis Janos Bolyai, 27 (1981), pp. 835–873.
- [19] G. SANGALLI, *Capturing small scales in elliptic problems using a residual-free bubbles finite element method*, Multiscale Model. Simul., 1 (2003), pp. 485–503.
- [20] R. E. SHOWALTER, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Math. Surveys Monogr. 49, American Mathematical Society, Providence, RI, 1997.
- [21] I. V. SKRYPNIK, *Methods for Analysis of Nonlinear Elliptic Boundary Value Problems*, Transl. Math. Monogr., American Mathematical Society, Providence, RI, 1994 (translated from the 1990 Russian original by Dan D. Pascali).
- [22] A. ŽAANEN, *An Introduction to the Theory of Integration*, North-Holland, Amsterdam, 1961.