View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Nottingham ePrints

# A Hierarchical Finite Element Monte Carlo Method for Stochastic Two-Scale Elliptic Equations

D.L. Brown*, V. H. Hoang†

April 6, 2017

**Abstract**

We consider two-scale elliptic equations whose coefficients are random. In particular, we study two cases: in the first case, the coefficients are obtained from an ergodic dynamical system acting on a probability space, and in the second the case, the coefficients are periodic in the microscale but are random. We suppose that the coefficients also depend on the macroscopic slow variables. While the effective coefficient of the ergodic homogenization problem is deterministic, to approximate it, it is necessary to solve cell equations in a large but finite size "truncated" cube and compute an approximated effective coefficient from the solution of this equation. This approximated effective coefficient is, however, realization dependent; and the deterministic effective coefficient of the homogenization problem can be approximated by taking its expectation. In the periodic random setting, the effective coefficient for each realization are obtained from the solutions of cell equations which are posed in the unit cube, but to compute its average by the Monte Carlo method, we need to consider many uncorrelated realizations to accurately approximate the average. Straightforward employment of finite element approximation and the Monte Carlo method to compute this expectation with the same level of finite element resolution and the same number of Monte Carlo samples at every macroscopic point is prohibitively expensive. We develop a hierarchical finite element Monte Carlo algorithm to approximate the effective coefficients at a dense hierarchical network of macroscopic points. The method requires an optimal level of complexity that is essentially equal to that for computing the effective coefficient at one macroscopic point, and achieves essentially the same accuracy. The levels of accuracy for solving cell problems and for the Monte Carlo sampling are chosen according to the level in the hierarchy that the macroscopic points belong to. Solutions and the effective coefficients at the points where the cell problems are solved with higher accuracy and the effective coefficients are approximated with a larger number of Monte Carlo samples are employed as correctors for the effective coefficient at those points at which the cell problems are solved with lower accuracy and fewer Monte Carlo samples. The method combines the hierarchical finite element method for solving cell problems at a dense network of macroscopic points with the optimal complexity developed in D. L. Brown, Y. Efendiev and V. H. Hoang, *Multiscale Model. Simul.* **11** (2013), with a hierarchical Monte Carlo sampling algorithm that uses different number of samples at different macroscopic points depending on the level in the hierarchy that the macroscopic points belong to. Proof of concept numerical examples confirm the theoretical results.

## 1 Introduction

In heterogeneous media, the primary challenge in obtaining accurate simulations that are cost effective is the multiscale nature of the media. This multiscale nature often necessitates the use of very fine scale meshes to resolve the characteristics of the heterogeneities. This difficulty is compounded by the fact that often the uncertainties in the material parameters may be large. This multiscale nature and uncertainty in material properties are exhibited in a variety of applications such as porous media, i.e. filtration or batteries devices for industrial applications [11, 25] or

*School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK (Donald.Brown@nottingham.ac.uk),

†Division of Mathematical Sciences, School of Physical & Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371 (vhhoang@ntu.edu.sg).

subsurface modeling in the context of groundwater modeling or oil and gas extraction [12, 14]. This combination of scales and the uncertainty yields problems that are not only high resolution, but may also require many realizations to handle the randomness.

In any setting, with these issues, new computational tools need to be developed to handle the multiple scales and randomness. The following is just a sample of the many tools and techniques that have been developed in this area. One such tool that has been successfully used to handle multiscale periodic coefficients as well as random media is homogenization. The theory of this is well established and the literature is quite extensive c.f. [6, 4, 5, 23, 26] to cite just a few. The primary idea is to derive an equation that only depends on the macroscopic scale of the domain and represents the multiscale equation macroscopically, i.e. the solution of the homogenized equation is equivalent to the solution of the multiscale equation in the average. To establish this equation, it is necessary to solve cell equations at the microscale that are posed in representative domain, and in the periodic setting, in the periodic cube. The equation is then upscaled or averaged to build a homogenized macroscale equation. In reservoir engineering literature this process is commonly known as upscaling and upgridding, where the calculation of fine-grid or subgrid is used to generate a coarser representation of the underlying geology [16, 21]. The computational complexity of upscaling, or of establishing the homogenized equation can be prohibitively high, as for each macroscopic points, several cell problems have to be solved; the only exception is when the microstructure is strictly periodic.

There have been attempts to solve the multiscale equations directly to obtain both the macroscopic and the microscopic information, without forming the homogenized, or effective, equations. The multiscale finite element methods (MsFEM) [13] solve local problems with fine mesh to build basis functions that contain microscopic information. The heterogeneous multiscale methods (HMM) [3] solve local homogenization cell problems at coarse-grid nodes to compute unknown coefficients. Other methods base on the variational multiscale principles ([20]) such as the local orthogonal decomposition (LOD) [10, 22] where a fine-scale space is constructed and sub-grid corrections are made to incorporate multiscale information. However, these methods do not utilize the periodic, almost periodic or ergodic structures of the media and may be computationally costly in the offline phase. We will focus ourselves in this work in utilizing special ergodic or microscale periodic structure of the random multiscale coefficients to develop a computationally efficient scheme for establishing the homogenized equation.

There are techniques to efficiently computationally homogenize multiscale elliptic equations. For periodic multiscale problems sparse tensor product finite elements can be used to solve the high dimensional multiscale homogenized problems to obtain all the necessary macroscopic and microscopic information, without forming the homogenized equation; the method achieves essentially optimal complexity [19]. Another approach, in the context of pore-scale fluid flows, utilizes a reduced-basis model reduction framework that relies on a parametrization of the microstructure [1]. For random multiscale elliptic equations, the authors in [18] approximate the random high dimensional coefficients by the best $N$ term approximations of the generalized polynomial chaos (GPC) expansions of the solution, combining with sparse tensor finite elements for the GPC coefficients of the scale interaction terms to obtain approximations with optimal complexity. Other works include utilization of the multilevel Monte Carlo technique by varying domain size [15]. In this paper, we develop an optimal method to compute the homogenized coefficient for two-scale ergodic and microscale periodic random problems for a dense hierarchical network of macroscopic points. For deterministic two-scale Stokes-Darcy systems, such an optimal method has been developed theoretically and verified numerically in the context of pore-scale fluid flows in deformable porous media [9]. The main idea of the algorithm is to solve cell problems at macroscopic points at different levels of the hierarchy using different levels of resolution. The method then uses a correction step that employs the solution at nearby macroscopic points at which cell problems are solved with higher accuracy to correct the lower resolution solutions. In this work, we develop an optimal method to compute homogenized coefficient of ergodic random and periodic random two-scale elliptic equations. The rigorous homogenization of such problems is well established [5, 7, 26].

At each macro point in the domain, in the ergodic setting, abstract cell problems involving the generator of the semigroup of the dynamical system in the probability space have to be solved. However, in reality, we need to choose a large representative volume in the physical variable

and solve a cell equation inside this volume. The homogenized coefficient is approximately the expectation of the effective coefficient obtained from the solution of this cell problems in the representative volume for all the realization. This expectation can be computed by the Monte Carlo method. In the microscale periodic random case, the effective coefficient is random and is obtained from the solutions to cell problems in the unit cube for each macro point and for each realization. We aim to compute the expectation of the effective coefficient at each macro point.

Thus, in general, there are three sources of errors that we need to account for. The first one is that of taking a finite size representative volume in which we consider the cell problems (in the ergodic case). The second error is that of solving these cell problems by using a discretization method; here we use a finite element method. The third error comes from Monte Carlo approximation of the effective coefficient where a finite number of samples is used. A straightforward combination of finite element solves of the cell problems and Monte Carlo approximations at every macroscopic point is prohibitively expensive, as we need to use a large number of samples for the Monte Carlo approximation, and for each sample, we need reasonably small meshes to solve the cell problems to get good accuracy. To expedite this computation, we will develop a hierarchy of grids of macro points and a corresponding nested collection of finite element (FE) approximation spaces. Further, we will utilize a hierarchical Monte Carlo Sampling technique.

The paper is organized as follows. In the next section, we formulate the stochastic elliptic equations and recall basic properties of ergodic homogenization for elliptic problems. We then present the error analysis for a straightforward combination of finite element approximation for cell problems in the truncated finite representative domain and Monte Carlo approximation. We also outline the differences in the microscale periodically random case and their effects on the error analysis. We then present the hierarchical finite element Monte Carlo (sometimes denoted MC) algorithm for two-scale stochastic homogenization in Section 3. Rigorous analysis on the error estimates of the hierarchical finite element method for solving cell problems for each fixed realization and the error estimates for the hierarchical Monte Carlo approximation using different number of samples at different macro points is then performed. Finally, in Section 4 we present some proof of concept numerical examples to verify the theoretical results. The paper ends with some conclusions in Section 5. The Appendix presents the lengthy proofs of some results before.

Throughout the paper, by $\nabla$ and by $\Delta$ we mean the gradient and the Laplacian respectively with respect to the variable $y$. The letter $c$ denotes various positive constants which do not depend on the approximation parameters such as the size $\rho$ of the representative domain, the mesh size and the number of Monte Carlo samples. Finally, $|\cdot|_{H^1(D)}$ and $|\cdot|_{H^2(D)}$ denote the seminorms in $H^1(D)$ and $H^2(D)$ respectively.

# 2 Problem Formulation

In this section, we present the general ergodic two-scale random coefficient case where we recall the derivation of the effective coefficients via solutions of the abstract cell problems. We discuss the combined errors for approximating the effective coefficients by finite element approximation and Monte Carlo sampling. We then summarize the periodic random two scale problems and highlight the difference in the methods for computing the average of the effective coefficients, and in the combined errors.

## 2.1 Multiscale Ergodic Random Elliptic Problems

Let $D$ be a domain in $\mathbb{R}^d$ where $d = 2, 3$. We consider a general elliptic problem with Dirichlet boundary condition on $D$

$$-\nabla \cdot (A^\varepsilon \nabla u^\varepsilon) = f \text{ in } D, \tag{1a}$$

$$u^\varepsilon = 0 \text{ on } \partial D, \tag{1b}$$

where $f$ is given data. Let $V = H_0^1(D)$ and $f \in V'$. The coefficient depends on the macroscopic scale of the domain $D$, and also on a microscopic scale $\varepsilon > 0$ which differs from the macroscopic scale by several degrees of magnitude. The coefficient is therefore denoted as $A^\varepsilon(x)$. In addition, we assume that the coefficient is random, and for now ergodic. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space.

For each $x$, we assume that $A^\varepsilon(x)$ is a random variable from $\Omega$ to $\mathbb{R}^{d\times d}_{sym}$, the space of symmetric matrices. We will focus on the following special ergodic structure of the the random multiscale coefficient. We assume that there is an ergodic dynamical system $T : \mathbb{R}^d \times \Omega \to \Omega$ such that for each $y, y' \in \mathbb{R}$, $T(y + y')\omega = T(y)T(y')\omega$ and every invariant subset of $\Omega$ under the map $T(y)$ has probability 0 or 1. Let $A$ be a function from $\mathbb{R} \times \Omega \to \mathbb{R}^{d\times d}_{sym}$ such that there are constants $c_1, c_2 > 0$ so that for all vectors $\xi, \zeta \in \mathbb{R}^d$,

$$c_1|\xi|^2 \le A(x,\omega)\xi \cdot \xi, \quad A(x,\omega)\xi \cdot \zeta \le c_2|\xi||\zeta|. \tag{2}$$

The two-scale random coefficient is defined as

$$A^\varepsilon(x,\omega) = A\left(x, T\left(\frac{x}{\varepsilon}\right)\omega\right). \tag{3}$$

For conciseness, for $x \in D$, $y \in \mathbb{R}^d$ and $\omega \in \Omega$, we denote by

$$A(x, y, \omega) = A(x, T(y)\omega).$$

To ease the presentation for showing that the coefficient in the $H^2(Y)$ regularity estimate (57) of $N^\rho$ is independent of $\rho$, in this paper, we assume that the coefficient is isotropic, and the matrix function $A(x,\omega)$ and $A^\varepsilon(x,\omega)$ are understood as a scalar function.

Homogenization of two-scale elliptic equations with ergodic stochastic coefficient (1) is well established (see e.g. [6], [24] and [26]). For example, the simplest ergodic structure is when $\Omega$ is the periodic torus $Y = [0,1]^d$ in $\mathbb{R}^d$ and $T(y)$ is the translation map. Using now classical two-scale asymptotic techniques [24], we write the solution to (1) as $u_\varepsilon$, we expand as

$$u^\varepsilon = u_0(x) + \varepsilon N_i(x,y)\frac{\partial u_0}{\partial x_i}(x) + \cdots,$$

where $y = \frac{x}{\varepsilon}$. The functions $N^i(x,y)$, $i = 1, \ldots, d$, which are periodic with respect to $y$ satisfy the cell problems

$$-\nabla \cdot \big(A(x,y)(\nabla N_i(x,y) + e^i)\big) = 0 \text{ in } Y, \tag{4}$$

where $e^i$ is the $i$-th unit vector in $\mathbb{R}^d$. We obtain the effective tensor given by

$$A^*_{ij}(x) = \frac{1}{|Y|}\int_Y A_{kj}(x,y)\left(\frac{\partial N_i}{\partial y_k}(x,y) + \delta_{ik}\right)dy.$$

For a general multiscale ergodic coefficient, to compute the effective coefficient $A^*$, for each unit vector $e^i \in \mathbb{R}^d$, we consider the abstract cell problem

$$\partial \cdot (A(x,\omega)(\partial N_i(x,\omega) + e^i)) = 0, \tag{5}$$

where $\partial = (\partial_1, \ldots, \partial_d)$ with $\partial_i$ being the generator of the semigroup $T(x + e_i)$. The coefficient $A^*$ is determined by

$$A^*_{ij}(x) = \int_\Omega A_{kj}(x,\omega)(\partial_k N_i(x,\omega) + \delta_{ik})d\mathbb{P}(\omega). \tag{6}$$

From Birkhoff theorem, the homogenized coefficient $A^*$ can also be determined by

$$A^*_{ij}(x) = \lim_{\rho\to\infty}\frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,T(y)\omega)\left(\frac{\partial N_i}{\partial y_k}(x,T(y)\omega) + \delta_{ik}\right)dy. \tag{7}$$

**Remark** We refer to [26] for the theoretical details. Though the authors in [26] only consider the case where the two-scale coefficient does not depend on the macroscopic variable $x$, i.e the function $A(x,\omega) = A(\omega)$, the proofs for the results for the case where $A$ depends on both $x$ and $\omega$ that we summarized above are exactly the same.

## 2.2 Finite Element Monte Carlo Approximation of Effective Coefficients

An analytic calculation of $A^*$ is in general impossible. However, the cell problem (5), and consequently the calculation of (7), can be reformulated in a finite domain with a Dirichlet boundary condition.

For the case where $A$ does not depend on $x$, i.e $A = A(\omega)$ and $A^\varepsilon(x) = A(T(\frac{x}{\varepsilon})\omega)$, the authors in [7] consider the cell problem in a truncated domain: find $N^\rho(y,\omega) \in H_0^1(Y^\rho)$ that satisfies

$$-\nabla_y \cdot \left(A(y,\omega)(\nabla_y N_i^\rho(y,\omega) + e^i)\right) = 0 \text{ in } Y^\rho, \tag{8a}$$

$$N_i^\rho(y,\omega) = 0 \text{ on } \partial Y^\rho, \tag{8b}$$

where $Y^\rho = [-\rho/2, \rho/2]^d$, or in $Y^\rho = [0,\rho]^d$ as in the numerical examples. We define the truncated approximation by

$$A_{ij}^{*,\rho}(\omega) = \frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(y,\omega)\left(\frac{\partial N_i^\rho}{\partial y_k}(y,\omega) + \delta_{ik}\right)dy. \tag{9}$$

Then for all $i,j = 1,\dots,d$

$$\lim_{\rho\to\infty} A_{ij}^{*\rho}(\omega) = A_{ij}^*$$

almost surely. Note that the truncated approximation effective coefficient $A^{*,\rho}$ is still random. The authors in [7] show the convergence estimate which depends on the size of the sample domain [7], and is given by

$$\mathbb{E}\left(\|A^* - A^{*,\rho}(\cdot)\|^2\right) \le \left(\frac{C}{\rho}\right)^\beta, \tag{10}$$

where the rate, $\beta > 0$, depends on the statistics of the media; $\mathbb{E}$ denotes the expectation with respect to the probability space $(\Omega, \Sigma, \mathbb{P})$, $\|\cdot\|$ denotes the norm in $\mathbb{R}^{d\times d}$ or the norm in $\mathbb{R}$ depending on the situation. We note that in (8) the Dirichlet boundary condition can be replaced by either the Neumann boundary condition or the periodic boundary condition.

In our current setting, we have the coefficient $A^\varepsilon(x,\omega) = A(x, T(\frac{x}{\varepsilon})\omega)$ so the truncated cell problem then becomes: find $N_i^\rho(x,y,\omega) \in H_0^1(Y^\rho)$ so that

$$-\nabla_y \cdot (A(x,y,\omega)(\nabla_y N_i^\rho(x,y,\omega) + e^i)) = 0 \text{ for } y \in Y^\rho, \tag{11a}$$

$$N_i^\rho(x,y,\omega) = 0, \text{ for } y \in \partial Y^\rho \tag{11b}$$

which in variational form becomes

$$\int_{Y^\rho} A(x,y,\omega)\nabla_y N_i^\rho(x,y,\omega)\cdot\nabla_y\phi(y)dy = -\int_{Y^\rho} A(x,y,\omega)e^i\cdot\nabla_y\phi(y)dy \quad \forall\,\phi \in H_0^1(Y^\rho).$$

The variational form of (11) is denoted

$$\mathcal{A}_y(x)\,(N_i^\rho(x,\cdot,\omega),\phi) = (F(x,\cdot,\omega),\phi), \tag{12}$$

for all $\phi \in H_0^1(Y^\rho)$, where $\mathcal{A}_y(x)\,(N_i^\rho(x,\cdot,\omega),\phi) = \int_{Y^\rho} A(x,y,\omega)\nabla_y N_i^\rho(x,y,\omega)\cdot\nabla_y\phi(y)dy$ and $(F(x,\cdot,\omega),\phi) = -\int_{Y^\rho} A(x,y,\omega)e^i\cdot\nabla_y\phi(y)dy$. We define the truncated effective coefficient as

$$A_{ij}^{*,\rho}(x,\omega) = \frac{1}{|Y|^\rho}\int_{Y^\rho} A_{kj}(x,y,\omega)\left(\frac{\partial}{\partial y_k}N_i^\rho(x,y,\omega) + \delta_{ik}\right)dy.$$

We make the following assumption.

**Assumption 2.1.** *We assume that there is a constant $C$, independent of $x$, so that for all $x \in D$*

$$\mathbb{E}\left(\|A^*(x,\omega) - A^{*,\rho}(x,\omega)\|^2\right) \le \frac{C}{\rho^\beta}.$$

From Assumption 2.1, for each $x \in D$, the effective coefficient $A^*(x)$ can be approximated by $\mathbb{E}(A^{*,\rho}(x,\omega))$. To compute this expectation, we develop the Monte Carlo algorithm, combining with finite element approximation for the truncated cell problem (11). For $M$ uncorrelated realizations $\{A(\cdot,\omega_1),\ldots,A(\cdot,\omega_M)\}$ of the media, we compute the solutions $\{N^\rho(x,\cdot,\omega_1),\ldots,N^\rho(x,\cdot,\omega_M)\}$ of problem (11). We denote the Monte Carlo average as

$$E^M(A_{ij}^{*,\rho}) = \frac{1}{M}\sum_{q=1}^{M} A_{ij}^{*,\rho}(\omega_q). \tag{13}$$

We have the following Monte Carlo estimate.

**Theorem 2.1.** *Under Assumption 2.1, for almost all $\omega \in \Omega$, the following convergence holds:*

$$\mathbb{E}(A_{ij}^{*,\rho}) = \lim_{M\to\infty} E^M(A_{ij}^{*,\rho}) = \lim_{M\to\infty}\frac{1}{M}\sum_{q=1}^{M} A_{ij}^{*,\rho}(\omega_q). \tag{14}$$

*Moreover, the root mean square error for the Monte Carlo approximation* (13) *to $A^*$ is*

$$\mathcal{E}^{\rho,M}(A_{ij}^*) = \sqrt{\mathbb{E}\left(|A_{ij}^* - E^M(A_{ij}^{*,\rho})|^2\right)} \le C\left(\left(\frac{1}{\rho}\right)^{\beta/2} + \left(\frac{1}{M}\right)^{\frac{1}{2}}\right) \tag{15}$$

*Proof.* See Appendix. □

The cell problem (11) in the truncated domain $Y^\rho$ needs to be solved approximately utilizing a finite element discretization and further introducing error. We partition $Y^\rho$ into a set of regular triangular simplices of the mesh size $h$. Let $V^h \subset H_0^1(Y^\rho)$ be the space of continuous functions which are linear in each simplex. The finite element problem is: find $N_i^{\rho,h} \in V^h$ so that

$$\int_{Y^\rho} A(x,y,\omega)(\nabla_y N_i^{\rho,h}(x,y,\omega) + e^i)\cdot\nabla_y\phi(y)dy = 0, \quad \forall\,\phi \in V^h. \tag{16}$$

We then have the following standard finite element error estimate where the constant $C_I$ only depends on the constants $c_1$ and $c_2$ in (2) and the shape of the simplices, and in particular, it is independent of the domain $Y^\rho$:

$$|N_i^\rho(x,\cdot,\omega) - N_i^{\rho,h}(x,\cdot,\omega)|_{H^1(Y^\rho)} \le C_I h |N_i^\rho(\cdot,\omega)|_{H^2(Y^\rho)}. \tag{17}$$

Following the notation above we denote the truncated finite element approximated effective coefficient by

$$A_{ij}^{*,\rho,h}(x,\omega) = \frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,y,\omega)\left(\frac{\partial N_i^{\rho,h}}{\partial y_k}(x,y,\omega) + \delta_{ik}\right)dy.$$

For $M$ uncorrelated samples $\omega_1,\omega_2,\ldots,\omega_M$, we denote Monte Carlo average $E^M(A_{ij}^{*,\rho,h})$ as

$$E^M(A_{ij}^{*,\rho,h}) = \frac{1}{M}\sum_{q=1}^{M} A_{ij}^{*,\rho,h}(\omega_q).$$

To bound $|N_i^\rho(x,\cdot,\omega)|_{H^2(Y^\rho)}$ for deriving the FE error in terms of the mesh size, we impose the following assumption on the coefficient $A$.

**Assumption 2.2.** *For $x \in D$, almost surely $|\partial_\omega A(x,\omega)| < c$ , with $c$ is independent of $x$ and $\omega$.*

We have the following result for the combined finite element and truncated error.

**Lemma 2.2.** *Under Assumptions 2.1 and 2.2, for all $i,j = 1,\ldots,d$*

$$\mathcal{E}^{\rho,M,h}(A_{ij}^*) = \sqrt{\mathbb{E}\left(|A_{ij}^* - E^M(A_{ij}^{*,\rho,h})|^2\right)} \le c\left(\left(\frac{1}{\rho}\right)^{\beta/2} + \left(\frac{1}{M}\right)^{\frac{1}{2}} + h\right), \tag{18}$$

*Proof.* See Appendix. □

### 2.3 Microscale Periodic Random Coefficients

We consider in this section two scale periodic random problems. Let $A$ be a function from $D\times Y\times\Omega$ to $\mathbb{R}^{d\times d}_{sym}$ so that for almost all $x\in D$ and almost all $\omega\in\Omega$, the function $A(x,\cdot,\omega)$ is continuous and periodic with the period being the unite cube $Y=(0,1)^d$ (Indeed, as mentioned above, to ease the presentation for showing that the coefficient in the $H^2(Y^\rho)$ regularity estimate (57) is independent of $\rho$, we will assume that $A$ is isotropic.) The two scale random coefficient is defined as

$$A^\varepsilon(x,\omega)=A\left(x,\frac{x}{\varepsilon},\omega\right).$$

We consider the two scale problem:

$$-\nabla\cdot(A^\varepsilon(x,\omega)\nabla u^\varepsilon)=f$$

with the Dirichlet boundary condition $u^\varepsilon(x)=0$ for $x\in\partial D$; the function $f\in V'$. For each realization, to determine the homogenized coefficient, we solve cell problems

$$-\nabla\cdot(A(x,y,\omega)(\nabla_y N_i(x,y)+e^i)=0$$

in the unit cube $Y$ where the function $N_i(x,y)$ is periodic with respect to $y$. The random effective coefficients are defined as

$$A^*_{ij}(x,\omega)=\int_Y A_{kj}(x,y,\omega)\left(\frac{\partial N_i}{\partial y_k}(x,y)+\delta_{ik}\right)dy.$$

We are interested in computing the expectation $\mathbb{E}(A^*(x,\omega))$ of the homogenized coefficient. This can be done by Monte Carlo method in an exactly the same way as above but the cell problems are now solved in the unit cube $Y$ (i.e. when $\rho=1$). Denoting the Monte Carlo average of the homogenized coefficient obtained from the finite element solution of the cell problems solved with mesh size $h$, using $M$ samples, as $E^M(A^{*,h})(x)$, we then have the error estimate:

**Proposition 2.1.** *For all $x\in D$, and all $i,j=1,\ldots,d$*

$$\mathbb{E}(|A^*_{ij}(x)-E^M(A^{*,h}_{ij})(x)|)^{1/2}\le c(M^{-1/2}+h).$$

*where $c$ is independent of $M$ and $h$.*

The proof follows exactly that for Lemma 2.2.

# 3 Hierarchical Finite Element Monte Carlo Algorithm

In the previous section, we approximated the effective coefficient by using Monte Carlo method to compute the expectation of the approximated coefficient on the truncated cell (or unit cell for the microscale periodic case) for each realization. For each macroscopic point $x\in D$, we need many samples for the approximation to be sufficiently accurate. Computing the effective coefficient for a large number of macroscopic points is thus very expensive. Adapting the approach for computing effective permeability of locally periodic microscopic structures developed by the authors in [9], we develop a hierarchical Monte Carlo algorithm to compute the effective permeability for the random multiscale coefficient in this section. We present the section for the ergodic random coefficients and will make remarks when there is difference in the algorithm for the periodic random case.

As in [9], we assume that the coefficient depends continuously on the macroscopic variable $x$. In particular, we assume the following Lipschitz bound

**Assumption 3.1.** *There is a constant $c$ such that for each $\omega\in\Omega$, for $x,x'\in D$*

$$\|A(x,\omega)-A(x',\omega)\|\le c|x-x'|. \tag{19}$$

We describe the hierarchical multiscale algorithm in the next section

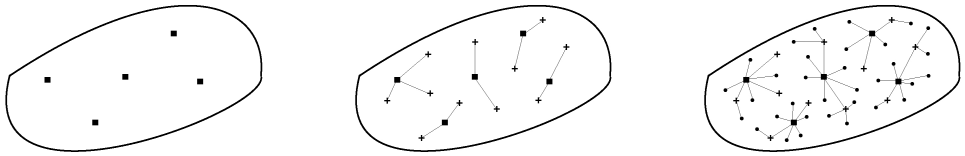

Figure 1: Schematic of one-point corrected macro-grid hierarchy of points.

### 3.1 Overview of Hierarchical Algorithm

We now develop a hierarchical method for solving the cell problem (11), and building an efficient approximation to $A^*(x)$ in (6). The algorithm outline follows closely the presentation in [9]. The first three steps are for computing the effective coefficient for a fixed realization, on a dense network of macroscopic points $x$ in the domain $D$. Due to the structure of the algorithm, we are able to utilize a Monte Carlo sampling procedure with different number of samples needed at different points in the hierarchically dense network of macro-grid points. This is done to expedite the computation of the MC average by requiring only few realizations at many points in the network. We outline the four step procedure below.

**Step 1 : Build Nested FE Spaces.** Fixing the macro-point $x \in D$ and a realization $\omega \in \Omega$, we wish to find an approximation for the solution $N_i^\rho(x, \cdot, \omega) \in H_0^1(Y^\rho)$, of (12), using Galerkin FEM. To this end, we build a nested collection of FE spaces. We denote the nested spaces as $\mathcal{V}_0 \subset \mathcal{V}_1 \subset \cdots \subset \mathcal{V}_L \subset H_0^1(Y^\rho)$ for a fixed positive integer $L$. These spaces are constructed so that the error of the Galerkin FE approximation (16) decreases in a structured way. More precisely, we assume the approximation properties

$$\inf_{\psi \in \mathcal{V}_l} \|v - \psi\|_{H_0^1(Y^\rho)} \leq C 2^{-l} h_0 |v|_{H^2(Y^\rho)} \tag{20}$$

for each $v \in H_0^1(Y^\rho) \bigcap H^2(Y^\rho)$ and $h_0$ being the coarsest mesh size corresponding to $\mathcal{V}_0$; the constant $C$ is independent of the cube $Y^\rho$ and only depends on the shape of the simplices. For example, we consider a hierarchy $\{\mathcal{S}_l\}_{l=1,2,\ldots}$ of triangular simplices in $Y^\rho$ which are constructed inductively by dividing each simplex in $\mathcal{S}_l$ into 4 congruent triangles when $d = 2$ and 8 tedrahedra when $d = 3$. The finite element space $\mathcal{V}_l$ consists of functions in $H_0^1(Y^\rho)$ which are linear in each simplex of $\mathcal{S}_l$.

**Step 2 : Build a Hierarchy of Macro-Grids.** We aim to solve the cell problem (12) at many macroscopic points $x \in D$ using different levels of resolution. We use the solution at points at which we solve this problem with high accuracy as the corrector for the solution at points at which we solve the problem with low accuracy. We will accomplish this by building a hierarchy of macro-grid points. First, we build a nested macro-grid for $D$ denoted by

$$\mathcal{T}_0 \subset \mathcal{T}_1 \subset \cdots \subset \mathcal{T}_L \subset D.$$

As in [9], we construct this grid inductively as follows. Suppose we have an initial grid $\mathcal{T}_0$ with grid spacing $H$ which is the maximal distance between neighboring nodes. Proceeding inductively, we construct refinement of $\mathcal{T}_{l-1}$, namely $\mathcal{T}_l$ with grid spacing $H2^{-l}$. The refinement of the macro-grid is related to the refinement factor of the error expression (20) for the nested FE spaces. Indeed, we see that formally we have error like $O((H2^{-l})(2^{-L+l}h_0)) = O(H2^{-L}h_0)$ or the same order as the finest approximation space $\mathcal{V}_L$.

We then define the dense hierarchy of macro-grids $\{S_0, S_1, \cdots, S_L\}$ inductively as $S_0 = \mathcal{T}_0$, $S_1 = \mathcal{T}_1 \backslash S_0$, and in general

$$S_l = \mathcal{T}_l \Big\backslash \left( \bigcup_{k<l} S_k \right).$$

We refer to the coarsest grid $S_0$ as the **anchor points**. We require that the hierarchy of macro-grids be **dense**. That is, we require that for each point $x \in S_l$, there exists at least one point $x' \in \bigcup_{k<l} S_k$ such that the $\text{dist}\,(x, x') < H2^{-l}$. A schematic example of a 3-level hierarchy of macro-grids $\{S_l\}_{l=0}^2$, contained in some domain can be seen in Figure 1, where we use a one-point correction to be discussed below.

**Step 3 : Calculating the Correction Term.**

To make an efficient multiscale method, we relate the nested FE spaces to the hierarchy of macro-grids. First, we solve cell problem (12) at the so-called anchor points, i.e. points at the most sparse macro-grid level $S_0$, with standard Galerkin FEM using $\mathcal{V}_L$, i.e. the highest level of accuracy. More specifically, for each unit vector $e^i \in \mathbb{R}^d$, we find $\bar{N}_i^\rho(x, \cdot, \omega) \in \mathcal{V}_L$ that satisfies

$$\mathcal{A}_y\,(x)\left(\bar{N}_i^\rho(x, \cdot, \omega), \phi\right) = \left(F(x, \cdot, \omega), \phi\right), \text{ for all } \phi \in \mathcal{V}_L. \tag{21}$$

We then proceed inductively to solve cell problem (12) for macroscopic points $x$ at other levels $S_l$ for $l = 1, \cdots, L$. Let $x \in S_l$ and $\{x_1, x_2, \cdots, x_n\} \subset (\bigcup_{k<l} S_k)$ be a collection of points such that the distance between $x$ and every point in the collection $\{x_1, x_2, \cdots, x_n\}$ is less than $O\left(H2^l\right)$ . Due to the density assumption in the construction of the hierarchy of grids, there exists at least one such point in $(\bigcup_{k<l} S_k)$. We denote the $l$-th macro-grid interpolation as

$$I_l^x(N_i^\rho) = \sum_{q=1}^{n} c_q N_i^\rho(x_q, \cdot, \omega), \tag{22}$$

where the constant weights $c_q$ satisfy $\sum_{q=1}^n c_q = 1$. Let $I_l^x(\bar{N}_i^\rho) = \sum_{q=1}^n c_q \bar{N}_i^\rho(x_q, \cdot, \omega)$ denote the macro-grid interpolation of Galerkin approximations. Recall, we assume that we have already computed the approximated solution $\{\bar{N}_i^\rho(x_q, \cdot, \omega)\}_{q=1}^n$. We solve for the correction term $\bar{N}_i^{\rho c}(x, \cdot, \omega) \in \mathcal{V}_{L-l}$ so that

$$\mathcal{A}_y\,(x)\left(\bar{N}_i^{\rho c}(x, \cdot, \omega), \phi\right) = \left(F(x, \cdot, \omega), \phi\right) - \mathcal{A}_y\,(x)\left(I_l^x(\bar{N}_i^\rho), \phi\right), \tag{23}$$

for all $\phi \in \mathcal{V}_{L-l}$. Note that the right-hand-side is known data from the previous finer accuracy solves at macro-grid points in $(\bigcup_{k<l} S_k)$. We solve for the correction term in the FE space $\mathcal{V}_{L-l}$ with coarser accuracy. Using both the correction term and the macro-grid interpolation term let

$$\bar{N}_i^\rho(x, \cdot, \omega) = \bar{N}_i^{\rho c}(x, \cdot, \omega) + I_l^x(\bar{N}_i^\rho), \tag{24}$$

be an approximation for $N_i^\rho(x, \cdot, \omega)$. We will show in Section 3.2 that the approximation (24), constructed via the hierarchical correction scheme is of the same accuracy order compared to standard way via (21) using the finest FE space $\mathcal{V}_L$, at an essentially optimal computational cost.

**Remark** As noted in [9], simplest macro-grid interpolation scheme is that of a single point, $I_l^x(v) = v(x_1, \cdot, \omega)$ for some $x_1 \in (\bigcup_{k<l} S_k) \bigcap B_{H2^l}(x)$. Here, $B_{H2^l}(x)$ is the open ball in $D$ centered at $x$ with radius $H2^l$. Another, being the two point scheme. For some $x_1, x_2 \in (\bigcup_{k<l} S_k) \bigcap B_{H2^l}(x)$ we write $I_l^x(v) = \frac{1}{2}(v(x_1, \cdot, \omega) + v(x_2, \cdot, \omega))$. We show the single point scheme in the schematic in Figure 1.

**Step 4 : Monte Carlo Sampling**

We now outline a MC sampling procedure that utilizes the hierarchical structure of the above. We proceed inductively. With the approximation $\bar{N}_i^\rho(x, \omega)$ of the solution to the cell problem (12) established above, we denote by $\bar{A}^{*,\rho}(x, \omega)$ the approximated effective coefficient by

$$\bar{A}_{ij}^{*,\rho}(x, \omega) = \frac{1}{|Y^\rho|} \int_{Y^\rho} A_{kj}(x, y, \omega) \left(\frac{\partial \bar{N}_i^\rho}{\partial y_k} + \delta_{ik}\right) dy.$$

For $x \in S_0$, i.e. $x$ is an anchor point, where we solve cell problem (12) with the highest resolution finite element space $\mathcal{V}_L$, we use $M_0 = O(2^{2L})$ samples to compute the MC average. Suppose that this procedure has been applied to find the Monte Carlo approximation for $\mathbb{E}(\bar{A}_{ij}^{*,\rho}(x'))$, which we

denote as $\mathbb{E}^{MC}(\bar{A}^{*,\rho}_{ij}(x'))$, and has been found for all $x' \in \mathcal{T}_{l'}$, for $l' < l$. We perform the MC sampling for approximating $\mathbb{E}(\bar{A}^{*,\rho}_{ij}(x))$ for $x \in S_l$ as follows: for

$$\bar{N}^{\rho}_i(x,\cdot,\omega) = \bar{N}^{\rho c}_i(x,\cdot,\omega) + \sum_{q=1}^{n} c_q \bar{N}^{\rho}_i(x_q,\cdot,\omega)$$

we define the quantities

$$I_1(x,\omega) = \sum_{q=1}^{n} c_q \frac{1}{|Y^\rho|} \int_{Y^\rho} (A_{kj}(x,y,\omega) - A_{kj}(x_q,y,\omega)) \left( \frac{\partial \bar{N}^{\rho}_i}{\partial y_k}(x_q,y,\omega) + \delta_{ik} \right) dy,$$

and

$$I_2(x,\omega) = \frac{1}{|Y^\rho|} \int_{Y^\rho} A_{kj}(x,y,\omega) \frac{\partial \bar{N}^{\rho c}_i}{\partial y_k}(x,y,\omega) dy.$$

We have

$$\bar{A}^{*,\rho}_{ij}(x,\omega) = \sum_{q=1}^{n} c_q \bar{A}^{*,\rho}(x_q,\omega) + I_1(x,\omega) + I_2(x,\omega). \tag{25}$$

We then use $M_l = O(2^{2(L-l)})$ samples to compute MC averages of $I_1(x,\cdot)$ and $I_2(x,\cdot)$. We let the following be an approximation for $\mathbb{E}(\bar{A}^{*,\rho}_{ij}(x,\cdot))$:

$$\mathbb{E}^{MC}(\bar{A}^{*,\rho}_{ij}(x,\cdot)) = \sum_{q=1}^{n} c_q \mathbb{E}^{MC}(\bar{A}^{*,\rho}_{ij}(x_q,\cdot)) + E^{M_l}(I_1(x,\cdot)) + E^{M_l}(I_2(x,\cdot)). \tag{26}$$

Thus, as we continue into the higher levels in the hierarchy, we require less MC samples.

## 3.2 Error Analysis

We will now show the error analysis for the hierarchical finite element and Monte Carlo algorithm in Section 3.1 for approximating the effective coefficient. We first need the following lemma on the Lipschitz dependence of the solution $N^\rho_i(x,y,\omega)$ of the truncated cell problem (12) on the macroscopic variable $x \in D$. Using the Lipschitz dependence of the coefficient in $x$ from Assumption 3.1, we then have the following result.

**Lemma 3.1.** *Under Assumptions 2.2 and 3.1, for $x, x' \in D$, and $s = 1, 2$, we have*

$$|N^\rho_i(x,\cdot,\omega) - N^\rho_i(x',\cdot,\omega)|_{H^s(Y^\rho)} \le c|x-x'||Y^\rho|^{1/2}.$$

*Proof* We first show the result for $H^1(Y^\rho)$ i.e.

$$\|\nabla N^\rho_i(x,\cdot,\omega) - \nabla N^\rho_i(x',\cdot,\omega)\|_{L^2(Y^\rho)} \le c|x-x'||Y^\rho|^{1/2}. \tag{27}$$

From the cell equations (11), for $x, x' \in D$, we have

$$\begin{aligned} -\nabla \cdot (A(x,y,\omega)\nabla(N^\rho_i(x,y,\omega) - N^\rho_i(x',y,\omega))) = \nabla \cdot ((A(x,y,\omega) - A(x',y,\omega))e^i) + \\ \nabla \cdot ((A(x,y,\omega) - A(x',y,\omega))\nabla N^\rho_i(x',y,\omega)). \end{aligned}$$

Therefore, using coercivity (2) and Lipschitz assumption (19) we obtain

$$\begin{aligned} &\|\nabla(N^\rho_i(x,\cdot,\omega) - N^\rho_i(x',\cdot,\omega))\|_{L^2(Y^\rho)} \\ &\quad \le \|A(x,\cdot,\omega) - A(x',\cdot,y)\|_{L^2(Y^\rho)} + \|A(x,\cdot,\omega) - A(x',\cdot,y)\|_{L^\infty(Y^\rho)} \|\nabla N^\rho_i(x',\cdot,\omega)\|_{L^2(Y^\rho)} \\ &\quad \le c|x-x'||Y^\rho|^{1/2} \end{aligned}$$

where we have used the inequality $\|\nabla_y N^\rho_i\|_{L^2(Y^\rho)} \le c|Y^\rho|^{1/2}$ shown in (51). From (11) we have

$$\Delta N^\rho_i(x,y,\omega) = A(x,y,\omega)^{-1}\nabla A(x,y,\omega) \cdot \nabla N^\rho_i(x,y,\omega) + A(x,y,\omega)^{-1}\nabla \cdot (A(x,y,\omega)e^i).$$

Therefore,

$$-\Delta(N_i^\rho(x,y,\omega)-N_i^\rho(x',y,\omega)) = A(x,y,\omega)^{-1}\nabla A(x,y,\omega)\cdot\nabla(N_i^\rho(x,y,\omega)-N_i^\rho(x',y,\omega)) + (A(x,y,\omega)^{-1}\nabla A(x,y,\omega)-A(x',y,\omega)^{-1}\nabla A(x',y,\omega))\cdot\nabla N^\rho(x',y,\omega) + (A(x,y,\omega)^{-1}\nabla\cdot(A(x,y,\omega)e^i)-A(x',y,\omega)^{-1}\nabla\cdot(A(x',y,\omega)e^i)).$$

We note that $\nabla A(x,y,\omega)=\partial_\omega A(x,T(y)\omega)$ and so is bounded. Thus the $L^2(Y^\rho)$ norm of the right hand side is not larger than $c|x-x'||Y^\rho|^{1/2}$. Using elliptic regularity, we deduce that

$$|N_i^\rho(x,\cdot,\omega)-N_i^\rho(x',\cdot,\omega)|_{H^2(Y^\rho)} \le c\|\nabla(N_i^\rho(x,\cdot,\omega)-N_i^\rho(x',\cdot,\omega))\|_{L^2(Y)} \le c|x-x'||Y^\rho|^{1/2}$$

where, by a simple rescaling argument, $c$ is independent of $\rho$, $x$ and $x'$. □

In the rest of this section, we show that the hierarchical finite element and Monte Carlo method we developed in the previous section to solve cell problem (11) and to approximate the effective coefficient achieves essentially optimal accuracy given the total number of degrees of freedom employed, similar to that for deterministic multiscale problems developed in [9]. To simplify the presentation, we show the error estimates for the case where the interpolation $I_l^x$ in (22) only uses one point $x'$. The general case is similar.

### 3.2.1 Error for fixed realization

First we show the error for approximating the solution $N_i^\rho(x,\cdot,\omega)$ of problem (12) by $\bar{N}_i^\rho(x,\cdot,\omega)$ constructed above, fixing a realization $\omega$. The proof is essentially similar to that for the deterministic problems considered in [9]. We need the estimates for the error analysis of the hierarchical Monte Carlo algorithm in the next subsection.

First we note that for each anchor point $x\in S_0$, (11) is solved by finite element using the finite element space $\mathcal{V}_L$ with the finite element solution $\bar{N}_i^\rho(x,\cdot,\omega)$ so the following error estimate holds

$$|N_i^\rho(x,\cdot,\omega)-\bar{N}_i^\rho(x,\cdot,\omega)|_{H^1(Y^\rho)} \le c2^{-L}|N_i^\rho(x,\cdot,\omega)|_{H^2(Y^\rho)} \le c2^{-L}|Y^\rho|^{1/2}$$

where we use the estimate (57). For a point $x\in S_l$, when $l>0$ we denote the one-point correction by

$$N_i^{\rho c}(x,y,\omega) = N_i^\rho(x,y,\omega)-N_i^\rho(x',y,\omega) \tag{28}$$

which satisfies the problem

$$\begin{aligned}&-\nabla\cdot(A(x,y,\omega)\nabla(N_i^{\rho c}(x,y,\omega))\\ &=\nabla\cdot(A(x,y,\omega)e^i)+\nabla\cdot(A(x,y,\omega)\nabla N_i^\rho(x',y,\omega)\\ &=\nabla\cdot((A(x,y,\omega)-A(x',y,\omega))e^i)+\nabla\cdot((A(x,y,\omega)-A(x',y,\omega))\nabla N_i^\rho(x',y,\omega))\end{aligned}$$

which can be written in terms of the variational form (12) as

$$\begin{aligned}\mathcal{A}_y(x)(N_i^{\rho c}(x,\cdot,\omega),\phi) &= (F(x,\cdot,\omega),\phi)-\mathcal{A}_y(x)(N_i^\rho(x',\cdot,\omega),\phi)\\ &= -\int_{Y^\rho}(A(x,y,\omega)-A(x',y,\omega))e^i\cdot\nabla\phi dy - \\ &\int_{Y^\rho}(A(x,y,\omega)-A(x',y,\omega))\nabla N_i^\rho(x',y,\omega)\cdot\nabla\phi dy\end{aligned} \tag{29}$$

for all $\phi\in H_0^1(Y^\rho)$. Problem (23) for one point corrector becomes

$$\int_{Y^\rho}A(x,y,\omega)\nabla\bar{N}_i^{\rho c}(x,y,\omega)\cdot\nabla\phi dy = -\int_{Y^\rho}(A(x,y,\omega)-A(x',y,\omega))e^i\cdot\nabla\phi dy - \int_{Y^\rho}(A(x,y,\omega)-A(x',y,\omega))\nabla\bar{N}_i^\rho(x',y,\omega)\cdot\nabla\phi dy,$$

for all $\phi\in\mathcal{V}_{L-l}$. The finite element approximation $\bar{N}_i^\rho(x,y,\omega)$ in the one point corrector case is

$$\bar{N}_i^\rho(x,y,\omega) = \bar{N}_i^\rho(x',y,\omega)+\bar{N}_i^{\rho c}(x,y,\omega). \tag{30}$$

We then have the following approximation:

**Proposition 3.1.** *For each point $x \in S_l$, there is a constant $c_l > 0$ such that*

$$|N_i^\rho(x,\cdot,\omega) - \bar{N}_i^\rho(x,\cdot,\omega)|_{H^1(Y^\rho)} \le c_l 2^{-L} h_0 |Y^\rho|^{1/2}.$$

*Proof.* From Lemma 3.1 we have that

$$|N_i^{\rho c}(x,\cdot,\omega)|_{H^2(Y^\rho)} \le c|x-x'||Y^\rho|^{1/2} \le c2^{-l}|Y^\rho|^{1/2}. \tag{31}$$

We prove the proposition by induction. Assume that the conclusion holds for all points in $\mathcal{T}_{l'}$ where $l' < l$. Let $\bar{\bar{N}}_i^{\rho c}(x,\cdot,\omega) \in \mathcal{V}_{L-l}$ be the solution of the problem:

$$\int_{Y^\rho} A(x,y,\omega)\nabla \bar{\bar{N}}_i^{\rho c}(x,y,\omega)\cdot\nabla\phi dy = -\int_{Y^\rho}(A(x,y,\omega)-A(x',y,\omega))e^i\cdot\nabla\phi dy - \int_{Y^\rho}(A(x,y,\omega)-A(x',y,\omega))\nabla N_i^\rho(x',y,\omega)\cdot\nabla\phi dy$$

for all $\phi \in \mathcal{V}_{L-l}$ which is the finite element approximation of (29). From Cea's lemma, the FE error corresponding to $\mathcal{V}_{L-l}$, and (31) we have

$$\|\nabla N_i^{\rho c} - \nabla \bar{\bar{N}}_i^{\rho c}\|_{L^2(Y^\rho)} \le c2^{-L+l}h_0|N^{\rho c}|_{H^2(Y^\rho)} \le c_1 2^{-L}h_0|Y^\rho|^{1/2} \tag{32}$$

where the constant $c_1$ is independent of $L$ and $\rho$. We then have

$$\int_{Y^\rho} A(x,y,\omega)\nabla(\bar{\bar{N}}_i^{\rho c}(x,y,\omega) - \bar{N}_i^{\rho c}(x,y,\omega))\cdot\nabla\phi dy = -\int_{Y^\rho}(A(x,y,\omega)-A(x',y,\omega))\nabla(N_i^\rho(x',y,\omega)-\bar{N}_i^\rho(x',y,\omega))\cdot\nabla\phi dy$$

for all $\phi \in \mathcal{V}_{L-l}$. Thus, using our induction hypothesis and the Lipschitz property of $A(\cdot,y,\omega)$, we obtain

$$\begin{aligned}
&\|\nabla(\bar{\bar{N}}_i^{\rho c}(x,y,\omega) - \bar{N}_i^{\rho c}(x,y,\omega))\|_{L^2(Y^\rho)} \\
&\quad\le \sup_{y\in Y^\rho}\|A(x,y,\omega)-A(x',y,\omega)\|\nabla(N_i^\rho(x',y,\omega)-\bar{N}_i^\rho(x',y,\omega))\|_{L^2(Y^\rho)} \\
&\quad\le c_2 2^{-l}c_{l'}2^{-L}h_0|Y^\rho|^{1/2}.
\end{aligned} \tag{33}$$

Therefore, we obtain

$$\begin{aligned}
&|N_i^\rho(x,\cdot,\omega) - \bar{N}_i^\rho(x,\cdot,\omega)|_{H^1(Y^\rho)} \\
&\quad= |N_i^\rho(x',\cdot,\omega) + N_i^{\rho c}(x,\cdot,\omega) - \bar{N}_i^\rho(x',\cdot,\omega) - \bar{N}_i^{\rho c}(x,\cdot,\omega)|_{H^1(Y^\rho)} \\
&\quad\le |N_i^\rho(x',\cdot,\omega) - \bar{N}_i^\rho(x',\cdot,\omega)|_{H^1(Y^\rho)} + |N_i^{\rho c}(x,\cdot,\omega) - \bar{\bar{N}}_i^{\rho c}(x,\cdot,\omega)|_{H^1(Y^\rho)} \\
&\qquad\quad + \|\bar{\bar{N}}_i^{\rho c}(x,\cdot,\omega) - \bar{N}_i^{\rho c}(x,\cdot,\omega)\|_{H^1(Y^\rho)} \\
&\quad\le c_{l'}2^{-L}h_0|Y^\rho|^{1/2} + c_1 2^{-L}h_0|Y^\rho|^{1/2} + c_2 c_{l'}2^{-l}2^{-L}h_0|Y^\rho|^{1/2}.
\end{aligned}$$

We then take

$$c_l = c_{l'} + c_1 + c_2 c_{l'}2^{-l}.$$

and arrive at our estimate. □

We then have the following approximation for the constant $c_l$.

**Proposition 3.2.** *There is a constant $c_*$ so that*

$$|N_i^\rho(x,\cdot,\omega) - \bar{N}_i^\rho(x,\cdot,\omega)|_{H^1(Y^\rho)} \le c_*(l+1)2^{-L}h_0|Y^\rho|^{1/2}$$

*when $x \in S_l$.*

*Proof.* We choose $\bar{l}$ independent of $L$ so that $l2^{-l} < 1/(2c_2)$ for $l > \bar{l}$. Let

$$c_* = \max\{\max_{0\le l\le \bar{l}} c_l/(l+1), 2c_1\}.$$

We prove by induction that for all $l$, $c_l \le c_*(l+1)$. This holds for $l \le \bar{l}$. For $l > \bar{l}$, assume that $c_{l'} < c_*(l'+1)$ for all $l' < l$. We have

$$\begin{aligned} c_l &\le c_*(l'+1) + c_1 + c_2 c_*(l'+1)2^{-l} \le c_*(l'+1) + c_*/2 + c_2 l 2^{-l} c_* \\ &\le c_*(l'+1) + c_*/2 + c_*/2 \le c_*(l'+2) \le c_*(l+1). \end{aligned}$$

□

### 3.2.2 Error with Monte Carlo Sampling

In Section 3.2.1 we present the error analysis for the hierarchical finite element algorithm for a single fixed realization to compute the effective coefficient. To approximate the effective coefficient $A^*(x)$, we need to compute the expectation of the approximated effective coefficient. The Monte Carlo method requires computing $\bar{A}^{*,\rho}(x,\omega)$ for many realizations $\omega$. A straightforward finite element approximation of the cell problems and Monte Carlo approximation of the expectation of the effective coefficient using fine mesh and full number of MC samples (which we refer to as the full reference solve in Section 4), and a straightforward Monte Carlo approximation that uses the hierarchical finite element algorithm to solve cell problems for each realization but uses the same large number of samples for each realization, (we refer to this as the hierarchical finite element approach in Section 4), has high complexity. We will show in this section that the combined hierarchical finite element Monte Carlo method presented in Section 3.1 that uses a number of samples depending on the level in the hierarchy that the macroscopic point $x$ belongs to achieves essentially optimal complexity. As in the previous subsection, to simplify the presentation, we show the error estimates for the case where there is only one point in the interpolation formula (22). In Section 4, we will refer to this technique as the combined hierarchical finite element Monte Carlo algorithm. We first recall that we defined the approximation to $A^{*,\rho}(x,\omega)$ as

$$\bar{A}^{*,\rho}_{ij}(x,\omega) = \frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,y,\omega)\left(\frac{\partial \bar{N}^\rho_i}{\partial y_k}(x,y,\omega) + \delta_{ik}\right) dy$$

where $\bar{N}^\rho_i(x,y,\omega)$ is solved using the hierarchical algorithm with grid and FE spaces outlined in Section 3.2.1. For $x \in S_l$ and for a fixed realization $\omega \in \Omega$ from Proposition 3.2 we have that

$$\begin{aligned} |A^{*,\rho}_{ij}(x,\omega) - \bar{A}^{*,\rho}_{ij}(x,\omega)| &= \Big|\frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,y,\omega)\left(\frac{\partial N^\rho_i}{\partial y_k}(x,y,\omega) - \frac{\partial \bar{N}^\rho_i}{\partial y_k}(x,y,\omega)\right) dy\Big| \\ &\le \frac{1}{|Y^\rho|}\|A(x,\cdot,\omega)\|_{L^2(Y^\rho)}\|\nabla N^\rho_i(x,\cdot,\omega) - \nabla \bar{N}^\rho_i(x,\cdot,\omega)\|_{L^2(Y^\rho)} \\ &\le \frac{1}{|Y^\rho|} c|Y^\rho|^{1/2} c_*(l+1)2^{-L}|Y^\rho|^{1/2} \\ &\le c c_*(l+1)2^{-L}. \end{aligned} \tag{34}$$

From (10) and (34),

$$\mathbb{E}\left(\|A^*(x) - \bar{A}^{*,\rho}(x,\omega)\|^2\right) \le c(\rho^{-\beta} + (l+1)^2 2^{-2L}),$$

so that

$$\|A^*(x) - \mathbb{E}(\bar{A}^{*,\rho}(x,\omega))\|^2 \le c(\rho^{-\beta} + (l+1)^2 2^{-2L}). \tag{35}$$

For $x, x' \in S = \cup_{l=1}^L S_l$, and using the FE approximation to (30), we obtain

$$\begin{aligned} \bar{A}^{*,\rho}_{ij}(x,\omega) - \bar{A}^{*\rho}_{ij}(x',\omega) &= \frac{1}{|Y^\rho|}\int_{Y^\rho}(A_{kj}(x,y,\omega) - A_{kj}(x',y,\omega))\left(\frac{\partial \bar{N}^\rho_i}{\partial y_k}(x',y,\omega) + \delta_{kj}\right) dy \\ &+ \frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,y,\omega)\frac{\partial \bar{N}^{\rho c}_i}{\partial y_k}(x,y,\omega) dy \\ &= I_1(x,\omega) + I_2(x,\omega) \end{aligned} \tag{36}$$

where we let

$$I_1(x,\omega) = \frac{1}{|Y^\rho|}\int_{Y^\rho}(A_{kj}(x,y,\omega) - A_{kj}(x',y,\omega))\left(\frac{\partial \bar{N}_i^\rho}{\partial y_k}(x',y,\omega) + \delta_{ik}\right)dy,$$

and

$$I_2(x,\omega) = \frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,y,\omega)\frac{\partial \bar{N}_i^{\rho c}}{\partial y_k}(x,y,\omega)dy.$$

For $x \in S_l$ and $x'$ is such that $|x - x'| \le c2^{-l}$ we have

$$\begin{aligned}|I_1(x,\omega)| &\le \frac{1}{|Y^\rho|}\left(\|A(x,\cdot,\omega) - A(x',\cdot,\omega)\|_{L^2(Y^\rho)}\|\nabla \bar{N}_i^\rho(x',\cdot,\omega)\|_{L^2(Y^\rho)}\right)\\ &\quad + \|A(x,\cdot,\omega)e - A(x',\cdot,\omega)e\|_{L^\infty(Y^\rho)}\\ &\le c|x-x'| \le c2^{-l}\end{aligned} \tag{37}$$

where $c$ is independent of $x$. Further, from Lemma 3.1, we have

$$|I_2(x,\omega)| \le \frac{1}{|Y^\rho|}\|A(x,\cdot,\omega)\|_{L^2(Y^\rho)}\|\nabla \bar{N}_i^{\rho c}(x,\cdot,\omega)\|_{L^2(Y^\rho)}. \tag{38}$$

Due to (27), (32) and (33), we deduce that $\|\nabla \bar{N}^{\rho c}\|_{L^2(Y^\rho)} \le c2^{-l}|Y^\rho|^{1/2}$. Therefore

$$|I_2(x,\omega)| \le c\frac{1}{|Y^\rho|}|Y^\rho|^{1/2}c2^{-l}|Y^\rho|^{1/2} \le c2^{-l}. \tag{39}$$

We have the following error bound for the hierarchical finite element Monte Carlo approximation of the effective coefficient.

**Theorem 3.2.** *Let $x \in S_l$, and let $\mathbb{E}^{MC}(\bar{A}^{*,\rho}(x,\cdot))$ in (26) be the Monte Carlo approximation of $\mathbb{E}(\bar{A}^{*,\rho}(x))$. Using $M_l = 2^{2(L-l)}$ samples to to compute $E^M(I_1(x,\cdot)), E^M(I_2(x,\cdot))$, we obtain the following error estimate*

$$\mathbb{E}\left(\|\mathbb{E}(A^*(x)) - \mathbb{E}^{MC}(\bar{A}^{*,\rho}(x))\|\right) \le C(\rho^{-\beta/2} + (l+1)2^{-L}) \tag{40}$$

*where $C$ is independent of $\rho$ and $L$.*

*Proof* For all $i,j = 1,\dots,d$, we first prove by induction that for a constant $C_*$ sufficiently large, for $x \in S_l$

$$\mathbb{E}(|\mathbb{E}(\bar{A}_{ij}^{*,\rho}(x,\omega)) - \mathbb{E}^{MC}(\bar{A}_{ij}^{*,\rho}(x))|) \le C_*(l+1)2^{-L}.$$

For points $x$ in level 0, with $M = O(2^{2L})$ Monte Carlo samples, we have

$$\mathbb{E}\left(|\mathbb{E}(\bar{A}_{ij}^{*,\rho}(x)) - \mathbb{E}^{MC}(\bar{A}_{ij}^{*,\rho}(x))|\right)^2 \le M^{-1}\mathbb{E}(|\mathbb{E}(\bar{A}_{ij}^{*,\rho}) - \bar{A}_{ij}^{*,\rho}|^2) \le c2^{-2L}$$

due to the fact that $|\bar{A}_{ij}^{*,\rho}|$ is uniformly bounded with respect to the realizations.

Assume that the Monte Carlo approximation for $\mathbb{E}(\bar{A}_{ij}^{*,\rho}(x'))$, which we denote as $\mathbb{E}^{MC}(\bar{A}^{*,\rho}(x'))$, has been found for all $x' \in \mathcal{T}_{l'}$, for $l' < l$ with

$$\left(\mathbb{E}\left(|\mathbb{E}(\bar{A}_{ij}^{*,\rho}(x',\omega)) - \mathbb{E}^{MC}(\bar{A}_{ij}^{*,\rho}(x'))|^2\right)\right)^{1/2} \le C_*(l'+1)2^{-L}. \tag{41}$$

As we use Monte Carlo sampling method with $M_l = 2^{2(L-l)}$ samples to approximate the expectation of $I_1(x,\omega)$, using estimate (37) we obtain

$$\mathbb{E}(|\mathbb{E}(I_1(x,\cdot)) - E^M(I_1(x,\cdot))|^2) \le C_1^2 2^{-2l}2^{-2(L-l)} \le C_1^2 2^{-2L} \tag{42}$$

for a positive constant $C_1$ independent of $x$, $l$ and $\rho$. Similarly, as we use $M_l = 2^{2(L-l)}$ Monte Carlo samples to approximate the expectation of $I_2(x,\omega)$, using (39) we obtain

$$\mathbb{E}(|\mathbb{E}(I_2(x,\cdot)) - E^M(I_2(x,\cdot))|^2) \le C_2^2 2^{-2l}2^{-2(L-l)} \le C_2^2 2^{-2L} \tag{43}$$

for a positive constant $C_2$ independent of $x$, $L$ and $\rho$. From (25), (26) for $n = 1$, and (41), (42), and (43) and have

$$\left(\mathbb{E}\left(|\mathbb{E}(\bar{A}_{ij}^{*,\rho}(x,\omega)) - \mathbb{E}^{MC}(\bar{A}_{ij}^{*,\rho}(x))|^2\right)\right)^{1/2} \leq C_*(l'+1)2^{-L} + C_1 2^{-L} + C_2 2^{-L}.$$

As the constants $C_1$ and $C_2$ are independent of $x$ and $L$, we choose $C_*$ such that $C_*/2 > C_1$ and $C_*/2 > C_2$. We then have

$$\mathbb{E}\left(\|\mathbb{E}(\bar{A}_{ij}^{*,\rho}(x)) - \mathbb{E}^{MC}(\bar{A}_{ij}^{*,\rho}(x))\|\right) \leq C_*(l'+2)2^{-L}) \leq C_*(l+1)2^{-L}. \tag{44}$$

We therefore have from (35)

$$\left(\mathbb{E}\left(|A_{ij}^{*}(x) - \mathbb{E}^{MC}(\bar{A}_{ij}^{*,\rho}(x))|\right)\right)^{1/2} \leq C(\rho^{-\beta/2} + (l+1)2^{-L}).$$

□

We assume that the hierarchy of macroscopic points $S_0, S_1, \ldots,$ are constructed so that the number of points in $S_l$ is $O(2^{dl})$. We then have the following results on the total degrees of freedom needed to perform the hierarchical finite element Monte Carlo algorithm.

**Proposition 3.3.** *For obtaining the homogenized coefficient for points in level $S_0, S_1, \ldots, S_L$ with the accuracy $O(L2^{-L})$ in the mean square, we need $O(2^{(d+2)L})$ degrees of freedom, which is optimal.*

*Proof* For each point in level $S_l$, we approximate $I_2$, for which we need to solve for $\bar{N}_i^{\rho c}(x,\cdot,\omega)$ with $2^{2(L-l)}$ samples. Thus the number of degrees of freedom required is

$$2^{2(L-l)}O(2^{d(L-l)}) = O(2^{(d+2)(L-l)}).$$

The number of points is $O(2^{dl})$ so that the total number of degrees of freedom required for each level $l$ is $O(2^{(d+2)(L-l)})O(2^{dl}) = O(2^{(d+2)L-2l})$. Thus the total number of degrees of freedom required is $O(2^{(d+2)L})$ which is asymptotically comparable to that required for computing the homogenized coefficient at one point with the same level of accuracy (using $2^{2L}$ Monte Carlo samples). □

**Remark** For periodic random two scale problems, we compute the expectation of the homogenized coefficient in an exactly the same way where the cell problems are solved inside the unit cube $Y$, i.e. the algorithm in Section 3.1 is applied verbatim where $Y^\rho$ is replaced by the unit cube $Y$. We drop the $\rho$ superscript in the notations for the finite element solution of the cell problems. From (44), we then have the error estimate for the hierarchical finite element Monte Carlo approximations for macroscopic points $x \in S_l$:

$$\mathbb{E}(\|\mathbb{E}(A^*(x)) - \mathbb{E}^{MC}(\bar{A}^*(x))\|) \leq C(l+1)2^{-L}.$$

# 4 Numerical Examples

In this section, we will apply our algorithm to two simple numerical examples, one for a two scale equation with an ergodic random coefficient, and one for an equation with a two scale periodic random coefficient. We will show that we obtain accurate results using the hierarchical finite element solve as described in [9] for the cell problems, and then use a large fixed number of Monte Carlo samples to approximate the effective coefficient for all macro points. We will refer to this approach as the hierarchical finite element method. We show also that the hierarchical finite element Monte Carlo approach described in Section 3.1 that chooses the level of finite element discretization and the number of Monte Carlo samples according to the levels that the macro points belong to give equally accurate approximations. We will compare these to what we call the full reference solve that uses the same fine mesh to solve the cell problems and the same large number of Monte Carlo samples to approximate the effective coefficients for all macro points. The reference solve will use at each macro-grid point the finest mesh discretization and the maximum number of Monte Carlo samples.

First, we will describe the macro-grid hierarchy $S_l$. We utilize the same grid as in [9] seen in Figure 2, where the lines indicate which points in the hierarchy are used to correct coarser solves.

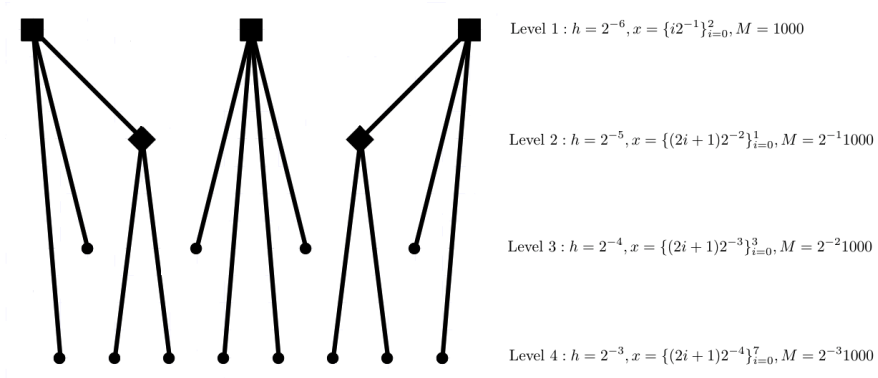


Figure 2: One dimensional macro-grid used in the numerical experiments. The lines indicate one-point correction locations. The squares indicate "anchor points", the diamonds indicates corrected points that again are used at the lower levels, and the dots utilizing the corrected solution at the higher levels.

We will suppose for these demonstrations of concept examples that the coefficients depend only on $x_1$ in the macroscopic variable and will depend on two-dimensional microscopic variables $y_1$ and $y_2$. We have four levels of macro-grids $S_l, l = 0, 1, 2, 3$. The level $S_0$ contains the so called anchor points at $x_1 = \{0, 1/2, 1\}$. The second level macro-grid points are given by $S_1 = \{1/4, 3/4\}$, and we note that in Figure 2, that the $*$ are used to signify that we will use the corrected solutions (corrected to full resolution) at these points further down in the hierarchy, for example, to correct solutions at $x_1 = \{3/16, 5/16\}$ and $x_1 = \{11/16, 13/16\}$, respectively. The third level, $S_2$, is given by $x_1 = \{1/8, 3/8, 5/8, 7/8\}$. Finally, the coarsest level, $S_3$, if given by $x_1 = \{(2k+1)/16\}, k = 0, \ldots, 7$.

For the finite element spaces $\mathcal{V}_{L-l}$, $l = 0, 1, 2, 3$, we use $\mathbb{P}_1$ Lagrange finite elements, on a regular triangular grid. As we must project from fine meshes to coarse mesh and prolongate from coarse meshes to fine mesh, using this regular triangular mesh is most suited as not to introduce interpolation errors. We will use standard projection and prolongation operators utilized in multigrid contexts [8]. We suppose that the finest scale space has mesh $h_f = 2^{-6}$. This corresponds to the sparsest macro-grid $S_0$. We then coarsen the space as $h_l = 2^{-6+l}$, so that the finite element space corresponding to $h_l$ is utilized in calculating corrections in macro-grid points $S_l$, for $l = 0, 1, 2, 3$. Finally, for the number of Monte Carlo samples, we will choose $\mathcal{O}(2^{2L})$ samples. Taking $L = 6$, the amounts to over 4000 samples. We will however take $M_0 = 1000 \propto \frac{1}{4}\mathcal{O}(2^{2L})$, as it is still proportional to our theoretical rate. Then, we take $M_1 = 500$, $M_2 = 250$, and $M_3 = 125$ samples when we combine both the spatially hierarchical solve and the hierarchical Monte Carlo sampling.

Due to computational software infrastructure considerations, we solve (8), by writing

$$\tilde{N}(x_1, y_1, y_2, \omega) = N(x_1, y_1, y_2, \omega) + y \cdot e,$$

and solve for $e = (1, 0)$ and $e = (0, 1)$

$$-\mathrm{div}_y \left( A(x_1, y_1, y_2, \omega) \nabla_y \tilde{N}(x_1, y_1, y_2, \omega) \right) = 0 \text{ in } Y, \tag{45a}$$

$$\tilde{N}(x_1, y_1, y_2, \omega) = y \cdot e \text{ on } \partial Y. \tag{45b}$$

This is merely done to make computing the permeability easier given our current computational infrastructure.

**Remark** Note in (45) we solve on the unit square $Y = [0, 1]^2$ as this is sufficient for periodic random coefficients. However, when we solve our ergodic quasi-periodic example in $Y^\rho = [0, \rho]^2$, we will merely map back the quantities to this unit cell and solve a $\rho$-dependent cell problem.

To compute the $A^*_{11}$ component of the effective coefficient we use

$$A^*_{11}(x_1,\omega) = \int_Y A(x_1,y_1,y_2,\omega)\nabla_y \tilde{N}(x_1,y_1,y_2,\omega)dy,$$
$$A^{*,ave}_{11,M}(x_1) = \frac{1}{M}\sum_{j=1}^{M} A^*_{11}(x_1,\omega_j). \tag{46}$$

Note, that similar cell problems are solved for $A^*_{ij}(x_1,\omega)$, for $i,j = 1,\ldots,d$.

## 4.1 Periodic Random Coefficient

As noted in the theory, the above approach of using hierarchical finite element for solving cell problems and hierarchical Monte Carlo method for approximating the average of the effective coefficient applies for both ergodic random and periodic random two scale coefficients. To test the algorithm on a simple, yet meaningful periodic random example we use the following coefficient example. This example is inspired by a structurally similar coefficient used by the authors in [2, Section 5] in the case of Multi-Level Monte Carlo methods. Suppose with a 1-dimensional macro-scale dependence and a two-dimensional micro-scale dependence we have

$$\begin{aligned} A(x_1,y,\omega) &= A_1(x_1,y_1)A_2(x_1,y_2)Y(\omega) \\ &= \left(1.1+\sin\left(\frac{\pi(x_1+.5)}{2}\right)+\sin(2\pi y_1)\right)\left(1.1+.5\sin\left(\frac{\pi(x_1+.5)}{2.4}\right)+\sin(2\pi y_2)\right)(.5\omega+1), \end{aligned} \tag{47}$$

for $\omega$, randomly independently distributed on $[0,1]$. Note that in the example the randomness can be factored out of the cell problem to calculate a deterministic numerical reference solution. This allows us to calculate the expectation of the random homogenized coefficient numerically by computing the effective coefficient with respect to the multiscale coefficient without the random part, and then multiplying with the expectation of the random variable, i.e.

$$A^{*,ave}(x_1) = \mathbb{E}(Y(\omega))A^*(x_1).$$

Recall, when computing the reference expectation of the homogenized coefficient, we solve the cell problems for the coefficient without the random part $Y(\omega)$ with the smallest mesh $h = 2^{-6}$, and then multiplying the numerical homogenized coefficient obtained with the expectation of the random variable $Y(\omega)$. The reference expectation of the homogenized coefficients are presented by the dash line – in Figure 3.

In the figure, the approximated average of the homogenized coefficient obtained from the hierarchical finite element algorithm for the cell problems, using the same large number of Monte Carlo samples at every macro point is shown in the o-points. We used 4 levels of macroscopic points and finite element approximations with mesh size $h_l = 2^{-6+l}, l = 0,1,2,3$, and the same number of Monte Carlo samples $M = 1000$ at every macro point. The *-points show the approximated average of the homogenized coefficient obtained from the hierarchical finite element Monte Carlo algorithm. We use the mesh size $h_l = 2^{-6+l}, l = 0,1,2,3$, to solve cell problems for macro points in level $l$ and use the number of Monte Carlo samples, $M_l = 1000, 500, 250, 125$, which are of order $\mathcal{O}(2^{2(6-l)})$ to approximate the expectation. We see that the algorithm is able to match very closely the reference effective coefficients. It is also the case that the hierarchical finite element Monte Carlo method performs remarkably similarly to the hierarchical finite element alone algorithm, but at much less cost in terms of number of samples needed. In Figure 4, we give a table of the results effective coefficient, (suppressing the subscript notation momentarily). We see that the relative errors tend to get larger as the corrected points are farther away from anchor points as well as with coarser finite element spaces.

## 4.2 Ergodic Quasi-Periodic Coefficient

Now we will consider a two scale equation with an ergodic coefficient. We choose to work with a quasi-periodic coefficient, where the ergodic randomness is introduced via phase shifts in the

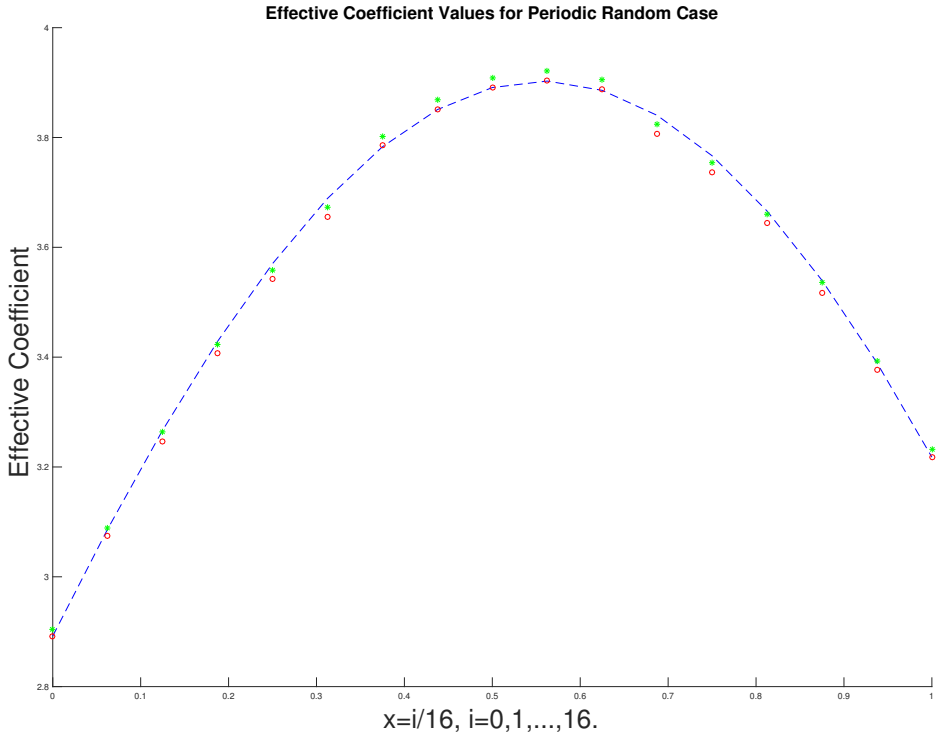


Figure 3: Using coefficient (47), $A_{11}^{*,ave}(x_1,\omega)$ is computed for $x_1 = i/16, i = 0,\ldots,16$. The reference solve is - - line, hierarchical solve is denoted in the o-points, and combined hierarchical solve in the *-points.

| $x_1$ | $A^*_{ref}(x_1)$ | $A^*_{hier}(x_1)$ | $A^*_{c-hier}(x_1)$ | %-$e_{hier}(x_1)$ | %-$e_{c-hier}(x_1)$ |
|---|---|---|---|---|---|
| 0 | 2.892034 | 2.892034 | 2.904612 | 0.0001 | 0.4349 |
| 1/16 | 3.087099 | 3.075311 | 3.089484 | 0.3818 | 0.0773 |
| 1/8 | 3.266906 | 3.246133 | 3.263167 | 0.6359 | 0.1145 |
| 3/16 | 3.428793 | 3.407658 | 3.422822 | 0.6164 | 0.1741 |
| 1/4 | 3.570252 | 3.542013 | 3.558347 | 0.7910 | 0.3334 |
| 5/16 | 3.689004 | 3.655897 | 3.673222 | 0.8975 | 0.4278 |
| 3/8 | 3.783064 | 3.785675 | 3.801272 | 0.0690 | 0.4813 |
| 7/16 | 3.850799 | 3.851499 | 3.868077 | 0.0182 | 0.4487 |
| 1/2 | 3.890973 | 3.890973 | 3.907895 | 0.0001 | 0.4349 |
| 9/16 | 3.902779 | 3.903438 | 3.920468 | 0.0169 | 0.4532 |
| 5/8 | 3.885866 | 3.888446 | 3.905337 | 0.0664 | 0.5011 |
| 11/16 | 3.840340 | 3.806224 | 3.824041 | 0.8884 | 0.4244 |
| 3/4 | 3.766769 | 3.737279 | 3.754496 | 0.7829 | 0.3258 |
| 13/16 | 3.666155 | 3.643793 | 3.660196 | 0.6100 | 0.1625 |
| 7/8 | 3.539909 | 3.517649 | 3.535416 | 0.6288 | 0.1269 |
| 15/16 | 3.389803 | 3.377010 | 3.392390 | 0.3774 | 0.0763 |
| 1 | 3.217906 | 3.217906 | 3.231901 | 0.0001 | 0.4349 |

Figure 4: Using coefficient (47), the effective coefficients of the reference, $A^*_{ref}(x_1)$, hierarchical, $A^*_{hier}(x_1)$, and combined hierarchical, $A^*_{c-hier}(x_1)$, algorithms. Along with percentage relative error from the reference value.

coefficient. First, we map the $\rho$-cell equation to a unit cell similar to (45). Here we map from the domain $Y^\rho = [0,\rho]^2$ to the unit domain $Y = [0,1]^2$. The mapped equation is given by

$$-\frac{1}{\rho^2}\mathrm{div}_y\left(A(x_1,\rho y_1,\rho y_2,\omega)\nabla_y\tilde{N}(x_1,\rho y_1,\rho y_2,\omega)\right) = 0 \text{ in } Y, \tag{48a}$$

$$\tilde{N}(x_1,\rho y_1,\rho y_2,\omega) = \rho y\cdot e \text{ on } \partial Y, \tag{48b}$$

and the mapped effective coefficient is given by

$$A_{11}^{*,\rho}(x_1,\omega) = \frac{1}{\rho}\int_Y A(x_1,\rho y_1,\rho y_2,\omega)\nabla_y\tilde{N}(x_1,\rho y_1,\rho y_2,\omega)dy.$$

For this example we will use the following (mapped) coefficient

$$\begin{aligned} &A(x_1,\rho y_1,\rho y_2,\omega_1,\omega_2)\\ &= 3.1 + (x_1+1)\sin(2\pi(\omega_1+\sqrt{2}(\rho y_1+\rho y_2))) + \sin(2\pi(\omega_2+\rho y_1+\rho y_2)), \end{aligned} \tag{49}$$

where $\omega = (\omega_1,\omega_2)\in[0,1]^2$. It is easy to verify that this quasi-periodic coefficient is ergodic.

Our purpose is to compare the effective coefficients obtained from the full solve and the hierarchical solve for one particular value of $\rho$ to demonstrate that the hierarchical solve obtained equally accurate approximations, but with much less cost. Recall, from the theory of [7], that we will have a theoretical error of the form $1/\rho^\beta$, where $\beta$ is often intractable and arises from the underlying randomness. In a numerical context, to gain some confidence of the convergence in $\rho$ we performed a simple study. This will inform us which $\rho$ value to fix for our computational tests. Using a sufficiently fine mesh, we compute $A_{11}^{*,\rho}(0,0)$, i.e., $x_1=\omega_1=\omega_2=0$, for varying values of $\rho$. We observe that as we take $\rho$ larger the value converged very quickly in this particular instance. We will thus take $\rho = 2^2$ as the solution appears to be sufficiently converged, but would not require too much resolution.

For this example, we utilized the same algorithm as in the periodic random coefficient. That is we had four macro-grid levels $S_l$ from Figure 2, the same resolution of finite element spaces, i.e. $h_l = 2^{-6+l}$, and the same levels of Monte Carlo samples $M_0 = 1000, M_1 = 500, M_2 = 250, M_3 = 125$. We take $\rho = 2^2$. The results of this simulation are recorded in Figure 5 as a plot and Figure 6 as a table with relative errors. We present for this test the reference permeability, $A^*_{ref}(x_1)$, using the finest mesh and full Monte Carlo sampling. We denote by $A^*_{hier}(x_1)$ the effective coefficient obtained from the finite element hierarchical solve with full Monte Carlo sampling, and by $A^*_{c-hier}(x_1)$ the effective coefficient obtained from the combined hierarchical finite element Monte Carlo algorithm in Section 3.1. Finally, we present the percentage relative error for both hierarchical solve, $\%$-$e_{hier}(x_1)$, and combined hierarchical solve $\%$-$e_{c-hier}(x_1)$. We note that all the results are within at most few percentage points away from the reference effective coefficient. We see that naturally, the anchor points in $S_0$ have the least error, while those points solved using coarser meshes introduce more error, but are computed at a much lower computational cost. We note that the maximal relative error is merely $\%3$.

In Figure 5 we plot these values as it is also worth noting that the highest errors are located along areas where the change in macro-scale features is the greatest. This corresponds to the variational form having a higher Lipschitz constant in the $x_1$ variable from Lemma 3.1. This is particularly pronounced at $x_1 = 11/16, 3/4$, where the highest percentage errors are observed. Finally we note that the combined hierarchical algorithm utilized less samples, however, the percentage errors are not greatly increased showing the computational effectiveness and efficiency of such an approach. We note the tendency for the values to be heavily influenced by the anchor points. The cause is two fold, the Lipschitz constant gets larger the farther away from the sample point, and the coarseness of the finite element space plays a role. One way to mitigate this tendency would be to use a correction that samples from the average solution of nearby points. That is, taking $\frac{1}{2}(\tilde{N}(0,y)+\tilde{N}(1/2,y))$, as the right hand side information to calculate the corrector in (23), for points in between $[0,1/2]$ for example. This suggests that in area where $A^*(x)$ changes the most requires a more refined macro-grid.

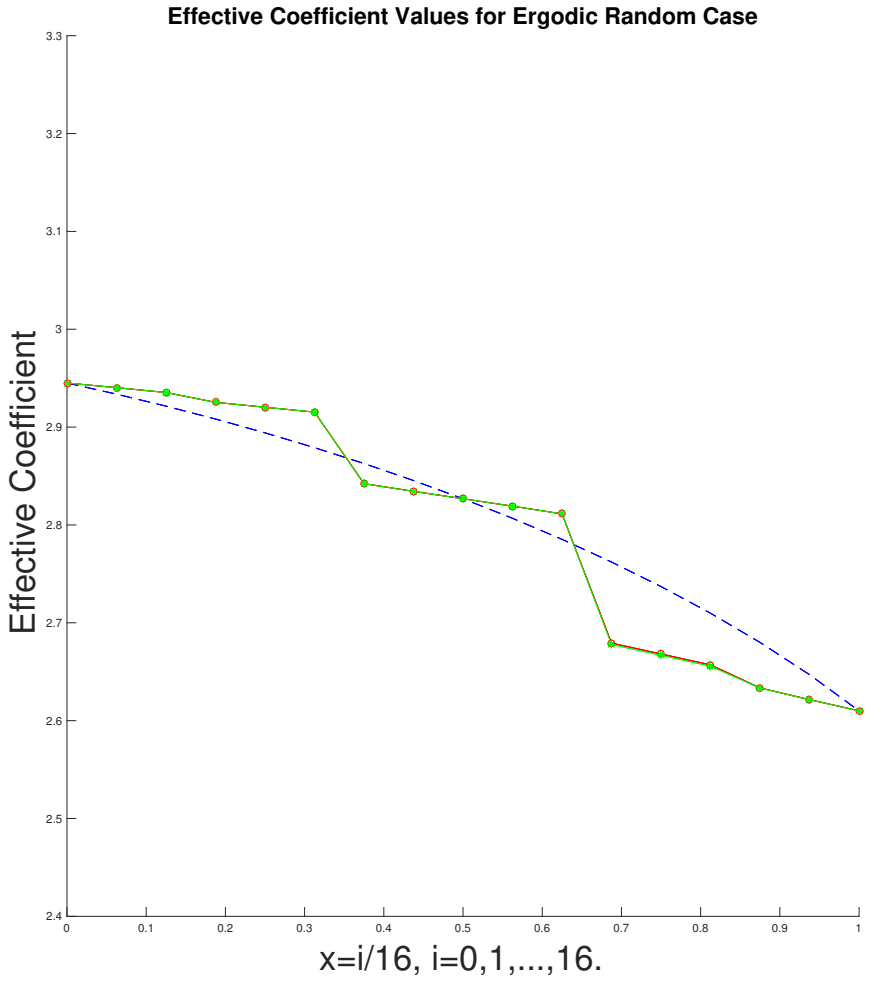


Figure 5: Using coefficient (49), $A_{11}^{*,ave}(x_1, \omega)$ is computed for $x_1 = i/16, i = 0, \ldots, 16$. The reference solve is - - line, hierarchical solve is denoted in the -o-points, and combined hierarchical solve in the -*-points.

# 5 Conclusions

In this work, we developed algorithms for the computationally demanding task of computing effective coefficients in random media with two scales. Along with developing the algorithm, precise error bounds were developed. This was done to show that by solving in some parts of the domain cell problems with greater resolution and Monte Carlo sampling, while using coarse solves and less samples in others, then using the fine-resolution solves nearby to correct, we are able to obtain an algorithm that is efficient as well as accurate. This was done for the case where the coefficients are periodic and random, yielding a non-deterministic effective coefficient as well as in an ergodic setting with a movable domain size. Finally, the algorithm was implemented on each case to demonstrate the applicability and main concepts. We showed good agreement between our algorithm and that of a reference solution using a full solve. Future work needs to be done to extend the method to more practical applications as well as parallelization of the macro-grid to utilize multiple processors. The ergodic example results in Figure 5 suggest that when designing the macro-grid, some consideration on the variation in the slow variable must be considered. This suggests a possible macro-grid adaptive strategy preprocessing strategy based on variations in the

| $x_1$ | $A^*_{ref}(x_1)$ | $A^*_{hier}(x_1)$ | $A^*_{c-hier}(x_1)$ | %-$e_{hier}(x_1)$ | %-$e_{c-hier}(x_1)$ |
|---|---|---|---|---|---|
| 0 | 2.945078 | 2.945075 | 2.945065 | 0.0001 | 0.0004 |
| 1/16 | 2.933687 | 2.940318 | 2.940325 | 0.2260 | 0.2263 |
| 1/8 | 2.921428 | 2.935345 | 2.935372 | 0.4764 | 0.4773 |
| 3/16 | 2.908261 | 2.925300 | 2.925305 | 0.5859 | 0.5861 |
| 1/4 | 2.894135 | 2.920301 | 2.920322 | 0.9041 | 0.9048 |
| 5/16 | 2.878996 | 2.915297 | 2.915334 | 1.2609 | 1.2622 |
| 3/8 | 2.862777 | 2.842215 | 2.842148 | 0.7183 | 0.7206 |
| 7/16 | 2.845400 | 2.834346 | 2.834299 | 0.3885 | 0.3901 |
| 1/2 | 2.826772 | 2.826775 | 2.826747 | 0.0001 | 0.0009 |
| 9/16 | 2.806779 | 2.819204 | 2.819195 | 0.4427 | 0.4424 |
| 5/8 | 2.785282 | 2.811291 | 2.811300 | 0.9338 | 0.9341 |
| 11/16 | 2.762107 | 2.679254 | 2.677826 | 2.9996 | 3.0513 |
| 3/4 | 2.737034 | 2.668133 | 2.666736 | 2.5174 | 2.5684 |
| 13/16 | 2.709772 | 2.657024 | 2.655661 | 1.9466 | 1.9969 |
| 7/8 | 2.679928 | 2.633427 | 2.633328 | 1.7352 | 1.7389 |
| 15/16 | 2.646952 | 2.621491 | 2.621403 | 0.9619 | 0.9652 |
| 1 | 2.610021 | 2.610039 | 2.609985 | 0.0007 | 0.0014 |

Figure 6: Using coefficient (49), the effective coefficients of the reference, $A^*_{ref}(x_1)$, hierarchical, $A^*_{hier}(x_1)$, and combined hierarchical, $A^*_{c-hier}(x_1)$, algorithms. Along with percentage relative error from the reference value.

coefficient on the large scale.

# 6 Appendix

Here we give the proofs of the Monte Carlo error estimates in Theorem 2.1 and in Lemma 2.2.

*Proof of Theorem 2.1.* This convergence is the consequence of the following estimate which shows the rate $M^{-1/2}$ for the convergence in the mean square.

$$\begin{aligned}
&\mathbb{E}(\|\mathbb{E}(A^{*,\rho}) - E^M(A^{*,\rho})\|^2) \\
&= \frac{1}{M^2}\sum_{k,l=1}^{d}\mathbb{E}\left(\sum_{i=1}^{M}\mathbb{E}(A^{*,\rho}_{kl}) - A^{*,\rho}_{kl}(\omega_i)\right)^2 \\
&= \frac{1}{M^2}\sum_{k,l=1}^{d}\mathbb{E}\left(\sum_{i=1}^{M}(\mathbb{E}(A^{*,\rho}_{kl}) - A^{*,\rho}_{kl}(\omega_i))^2 + 2\sum_{\substack{i,j=1\\ i\neq j}}^{M}(\mathbb{E}(A^{*,\rho}_{kl}) - A^{*,\rho}_{kl}(\omega_i))(\mathbb{E}(A^{*,\rho}_{kl}) - A^{*,\rho}_{kl}(\omega_j))\right) \\
&= \frac{1}{M^2}\sum_{k,l=1}^{d}\mathbb{E}\sum_{i=1}^{M}(\mathbb{E}(A^{*,\rho}_{kl}) - A^{*,\rho}_{kl}(\omega_i))^2 \\
&= \frac{1}{M}\mathbb{E}(\|\mathbb{E}(A^{*,\rho}) - A^{*,\rho}\|^2).
\end{aligned}$$

From this and (10) we have the estimate for the rate of convergence of $E^M(A^{*,\rho})$ to $A^*$:

$$\begin{aligned}
\mathbb{E}\left(\left\|A^* - E^M(A^{*,\rho})\right\|^2\right) &\leq 2\left\|A^* - \mathbb{E}(A^{*,\rho})\right\|^2 + 2\mathbb{E}\left(\left\|\mathbb{E}(A^{*,\rho}) - E^M(A^{*,\rho})\right\|^2\right) \\
&\leq C\left(\frac{1}{\rho}\right)^{\beta} + \frac{2}{M}\mathbb{E}\left(\left\|A^{*,\rho} - \mathbb{E}(A^{*,\rho})\right\|^2\right) \qquad (50)
\end{aligned}$$

where the constant $C$ is independent of $\rho$.

Next we will show that $\mathbb{E}\left(\|A^{*,\rho} - \mathbb{E}(A^{*,\rho})\|^2\right)$ is bounded above uniformly for all $\rho$. From (8) we have

$$\int_{Y^\rho} A(y,\omega)|\nabla N_i^\rho(y,\omega)|^2 dy = -\int_{Y^\rho} (A(y,\omega)e^i)\nabla N_i^\rho(y,\omega)dy.$$

Using the boundedness and coercivity conditions (2), for all $\omega \in \Omega$, we have by an application of the Cauchy Schwarz inequality

$$\|\nabla N_i^\rho(y,\omega)\|^2_{L^2(Y^\rho)} \leq C \|A(y,\omega)e\|_{L^2(Y^\rho)} \|\nabla N_i^\rho(y,\omega)\|_{L^2(Y^\rho)} \leq C|Y^\rho|^{\frac{1}{2}} \|\nabla N_i^\rho(y,\omega)\|_{L^2(Y^\rho)},$$

where $C$ depends on $c_1$ and $c_2$. Thus,

$$|N_i^\rho(y,\omega)|_{H^1(Y^\rho)} \leq C|Y^\rho|^{\frac{1}{2}}, \tag{51}$$

From the (9) and (51) we have

$$\begin{aligned}
|A_{ij}^{*,\rho}(\omega)| &\leq \frac{1}{|Y^\rho|}\left(\int_{Y^\rho} |A(y,\omega)\nabla_y N_i^\rho(y,\omega)|dy\right) + \frac{1}{|Y^\rho|}\left(\int_{Y^\rho} |A(y,\omega)e^i|dy\right) \\
&\leq \frac{C}{|Y^\rho|}\|A(\cdot,\omega)\|_{L^2(Y^\rho)} \|\nabla N_i^\rho(y,\omega)\|_{L^2(Y^\rho)} + \frac{c_2}{|Y^\rho|}\left(\int_{Y^\rho} |e^i|dy\right) \leq C
\end{aligned}$$

so $A^{*,\rho}(\omega)$ is uniformly bounded above for all $\rho$ and all $\omega \in \Omega$. Thus, we have

$$\mathbb{E}\left(\|A^{*,\rho} - \mathbb{E}(A^{*,\rho})\|^2\right) \leq C. \tag{52}$$

Thus, the root mean square error for estimating the effective permeability $A^*$ by Monte Carlo method is

$$\begin{aligned}
\mathcal{E}^{\rho,M}(A^*) = \sqrt{\mathbb{E}\left(\|A^* - E^M(A^{*,\rho})\|^2\right)} &\leq \left(\frac{C}{\rho}\right)^{\beta/2} + \left(\frac{C}{M}\right)^{\frac{1}{2}} \\
&\leq C\left(\left(\frac{1}{\rho}\right)^{\beta/2} + \left(\frac{1}{M}\right)^{\frac{1}{2}}\right).
\end{aligned} \tag{53}$$

Thus, we have proven our estimate. □

*Proof of Lemma 2.2.* To estimate this bound, we first establish the dependency of $|N^\rho(x,\cdot,\omega)|_{H^2(Y^\rho)}$ on $\rho$. From (8) we have

$$-\Delta N_i^\rho(x,y,\omega) = A^\rho(y,\omega)^{-1}(\nabla A^\rho(x,y,\omega)\nabla N_i^\rho(x,y,\omega) + \nabla \cdot \left(A^\rho(x,y,\omega)e^i\right)) \tag{54}$$

Using boundedness and coercivity (2) together with the condition $\nabla A(x,y,\omega) \in L^\infty(\mathbb{R}^d)$ uniformly for all $\omega \in \Omega$, and (51) we have

$$\|\Delta N_i^\rho(x,\cdot,\omega)\|_{L^2(Y^\rho)} \leq \|\nabla A^\rho(x,\cdot,\omega)\nabla N_i^\rho(x,\cdot,\omega)\|_{L^2(Y^\rho)} + \left\|\nabla \cdot \left(A^\rho(x,\cdot,\omega)e^i\right)\right\|_{L^2(Y^\rho)} \leq C|Y^\rho|^{\frac{1}{2}}. \tag{55}$$

Form elliptic regularity (see, e.g. [17] Theorem 3.2.1.3) we have

$$|N_i^\rho(x,y,\omega)|_{H^2(Y^\rho)} \leq C\|\Delta N_i^\rho(x,y,\omega)\|_{L^2(Y^\rho)}, \tag{56}$$

where, by a simple scaling argument, the constant $C$ above does not depend on $\rho$. Thus,

$$|N_i^\rho(x,y,\omega)|_{H^2(Y^\rho)} \leq C|Y^\rho|^{\frac{1}{2}} \tag{57}$$

where $C$ is independent of $\rho$. Therefore

$$|N_i^\rho(x,\cdot,\omega) - N_i^{\rho,h}(x,\cdot,\omega)|_{H^1(Y^\rho)} \leq Ch|Y^\rho|^{1/2}.$$

Recall, the approximated effective permeability obtained from the finite element solution $N_i^{\rho,h}$ is

$$A_{ij}^{*,\rho,h} = \frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,y,\omega)\left(\frac{\partial N_i^{\rho,h}}{\partial y_k}(x,y,\omega) + \delta_{ik}\right) dy. \tag{58}$$

From (9) and (58) we have

$$\begin{aligned}\left\|A_{ij}^{*,\rho} - A_{ij}^{*,\rho,h}\right\| &= \left|\frac{1}{|Y^\rho|}\int_{Y^\rho} A_{kj}(x,y,\omega)\left(\frac{\partial N^\rho}{\partial y_k}(x,y,\omega) - \frac{\partial N^{\rho,h}}{\partial y_k}(x,y,\omega)\right) dy\right| \\ &\leq \frac{1}{|Y^\rho|}\left\|A(x,\cdot,\omega)\right\|_{L^2(Y^\rho)}\left\|\nabla N_i^\rho(x,y,\omega) - \nabla N_i^{\rho,h}(x,y,\omega)\right\|_{L^2(Y^\rho)} \\ &\leq \frac{1}{|Y^\rho|}C|Y^\rho|^{1/2}Ch|Y^\rho|^{1/2} \qquad (59) \\ &\leq Ch \qquad (60)\end{aligned}$$

where the constant $C$ is independent of $\rho$ and of the mesh-size $h$. From this we have

$$\mathbb{E}(\|A^* - A^{*,\rho,h}\|^2) \leq c(\rho^{-\beta} + ch^2). \tag{61}$$

For using the Monte Carlo algorithm to estimate the expectation, we use the approximated effective permeability $A^{*,\rho,h}$ obtained from the finite element solution of the cell problem (8). We denote $E^M(A^{*,\rho,h})$ by

$$E^M(A^{*,\rho,h}) = \frac{1}{M}\sum_{q=1}^{M} A^{*,\rho,h}(\omega_q).$$

We have

$$\mathbb{E}(\|\mathbb{E}(A^{*,\rho,h} - E^M(A^{*,\rho,h})\|^2) \leq \frac{1}{M}\mathbb{E}(\|\mathbb{E}(A^{*,\rho,h} - A^{*,\rho,h}\|^2) \leq cM^{-1}. \tag{62}$$

From (61) and (62) we have the root mean square error of the finite element Monte Carlo

$$\begin{aligned}\mathcal{E}^{\rho,M,h}(A^*) = \sqrt{\mathbb{E}\left(\|A^* - E^M(A^{*,\rho,h})\|^2\right)} &\leq \left(\frac{c}{\rho}\right)^{\beta/2} + \left(\frac{c}{M}\right)^{\frac{1}{2}} + ch \qquad (63) \\ &\leq c\left(\left(\frac{1}{\rho}\right)^{\beta/2} + \left(\frac{1}{M}\right)^{\frac{1}{2}} + h\right), \qquad (64)\end{aligned}$$

where the constant $C$ is independent of $\rho$, $M$ and $h$. □

**Acknowledgment** V. H. Hoang gratefully acknowledges the financial support of the AcRF Tier 1 grant RG30/16, the Singapore A*Star SERC grant 122-PSF-0007 and the AcRF Tier 2 grant MOE 2013-T2-1-095 ARC 44/13.

# References

[1] A. Abdulle and O. Budac. A petrov-galerkin reduced basis approximation of the stokes equation in parametrized geometries. *(submitted to C. R. Math. Acad. Sci. Paris)*, 2015.

[2] Assyr Abdulle, Andrea Barth, and Christoph Schwab. Multilevel monte carlo methods for stochastic elliptic multiscale pdes. *Multiscale Modeling & Simulation*, 11(4):1033–1070, 2013.

[3] Assyr Abdulle, Weinan E, Björn Engquist, and Eric Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012.

[4] A. Allaire. Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23:1482–1518, 1992.

[5] A. Yu. Beliaev and S. M. Kozlov. Darcy equation for random porous media. *Comm. Pure Appl. Math.*, 49(1):1–34, 1996.

[6] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*, volume 5 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1978.

[7] A. Bourgeat and A. Piatnitski. Approximations of effective coefficients in stochastic homogenization. In *Annales de l'IHP Probabilités et statistiques*, volume 40, pages 153–165, 2004.

[8] James H Bramble. *Multigrid methods*, volume 294. CRC Press, 1993.

[9] D. L. Brown, Y. Efendiev, and V. H. Hoang. An efficient hierarchical multiscale finite element method for stokes equations in slowly varying media. *Multiscale Modeling & Simulation*, 11(1):30–58, 2013.

[10] D.L. Brown, D. Gallistl, and D. Peterseim. Multiscale petrov-galerkin method for high-frequency heterogeneous helmholtz equations. *in press, LNCSE, arXiv:1511.09244*, 2015.

[11] D.L. Brown and V. Taralova. A multiscale finite element method for neumann problems in porous microstructures. *Disc. and Cont. Dyn. Sys Series S.*, 9((5)):1299–1326,, 2016.

[12] P. Dostert, Y. Efendiev, T. Hou, and W. Luo. Coarse-gradient Langevin algorithms for dynamic data integration and uncertainty quantification. *Journal of Computational Physics*, 217(1):123–142, 2006.

[13] Y. Efendiev and T. Hou. *Multiscale Finite Element Methods: Theory and Applications*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, 233 Spring Street, New York, NY 10013, USA, 2009.

[14] Y. Efendiev, T. Hou, and W. Luo. Preconditioning of MCMC simulations using coarse-scale models. *SIAM. Sci. Comp.*, 28(2):776–803, 2006.

[15] Y. Efendiev, C. Kronsbein, and F. Legoll. Multilevel monte carlo approaches for numerical homogenization. *Multiscale Modeling & Simulation*, 13(4):1107–1135, 2015.

[16] C. L. Farmer. Upscaling: a review. *International Journal for Numerical Methods in Fluids*, 40(1-2):63–78, 2002.

[17] Pierre Grisvard. *Elliptic problems in nonsmooth domains*, volume 69. SIAM, 2011.

[18] V. H. Hoang and Ch. Schwab. Analytic regularity and polynomial approximation of stochastic, parametric elliptic multiscale pdes. *Analysis and Applications*, 11:1350001, 2013.

[19] V.H. Hoang and C. Schwab. High-dimensional finite elements for elliptic problems with multiple scales. *Multiscale Model. Simul.*, 3:168–194, 2005.

[20] Thomas J. R. Hughes. Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127(1-4):387–401, 1995.

[21] Durlofsky L. J. Upscaling and gridding of fine scale geological models for flow simulation. In *8th International Forum on Reservoir Simulation Iles Borromees, Stresa, Italy*, volume 2024, 2005.

[22] Axel Målqvist and Daniel Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014.

[23] G. Nguetseng. A general convergence result for a functional related to the theory of homogenization. *SIAM J. Math. Anal.*, 20:608–623, 1989.

[24] E. Sanchez-Palencia. *Non-Homogeneous Media and Vibration Theory*, volume 127 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 1980.

[25] V. Taralova. *Upscaling approaches for nonlinear processes in lithium-ion batteries*. PhD thesis, TU Kaiserslautern, Germany, 2015.

[26] V.V Zhikov, S.M. Kozlov, and O.A. Oleinik. *Homogenization of Differential Operators and Integral Functionals*. Springer-Verlag, Berlin, 1994.