Multiscale Model. Simul.
Vol. 9, No. 4, pp. 1549–1587

CORRECTOR THEORY FOR MSFEM AND HMM IN RANDOM MEDIA*

GUILLAUME BAL† and WENJIA JING‡

Abstract. We analyze the random fluctuations of several multiscale algorithms, such as the multiscale finite element method (MsFEM) and the finite element heterogeneous multiscale method (HMM), that have been developed to solve partial differential equations with highly heterogeneous coefficients. Such multiscale algorithms are often shown to correctly capture the homogenization limit when the highly oscillatory random medium is stationary and ergodic. This paper is concerned with the random fluctuations of the solution about the deterministic homogenization limit. We consider the simplified setting of the one-dimensional elliptic equation, where the theory of random fluctuations is well understood. We develop a fluctuation theory for the multiscale algorithms in the presence of random environments with short-range and long-range correlations. For a given mesh size h, we show that the fluctuations converge in distribution in the space of continuous paths to Gaussian processes as the correlation length $\varepsilon \to 0$. We next derive the limit of such Gaussian processes as $h \to 0$ and compare this limit with the distribution of the random fluctuations of the continuous model. When such limits agree, we conclude that the multiscale algorithm captures the random fluctuations accurately and passes the corrector test. This property serves as an interesting benchmark to assess the behavior of the multiscale algorithm in practical situations where the assumptions necessary for the theory of homogenization are not met. What we find is that the computationally more expensive methods MsFEM, and HMM with a choice of parameter $\delta = h$, correctly capture the random fluctuations both for short-range and long-range oscillations in the medium. The less expensive method HMM with $\delta < h$ correctly captures the fluctuations for long-range oscillations and strongly amplifies their size in media with short-range oscillations. We present a modified scheme with an intermediate computational cost that captures the random fluctuations in all cases.

Key words. equations with random coefficients, multiscale finite element method, heterogeneous multiscale method, corrector test, correlation ranges

AMS subject classifications. 65C99, 65Y20, 34E13

DOI. 10.1137/100815918

1. Introduction.

Differential equations with highly oscillatory coefficients arise naturally in many areas of applied sciences and engineering from the analysis of composite materials to the modeling of geological basins. Often, it is impossible to solve the microscopic equations exactly, in which case we can either solve macroscopic models when they exist or devise multiscale algorithms that aim to capture as much of the microscopic scale as possible. When the coefficients are periodic functions or stationary and ergodic processes, the solution to the heterogeneous equation is often well approximated by the deterministic solution of a homogenized equation. This is the well-known *homogenization theory*. In many applications, such as parameter estimation and uncertainty quantification, estimating the random fluctuations (finding the random corrector) in the solution is as important as finding its homogenized limit.

Finding the homogenized coefficients may be a daunting computational task, and the assumptions necessary to the applicability of homogenization theory may not be met [18], [20], [22]. Several numerical methodologies have been developed to find accurate

*Received by the editors November 24, 2010; accepted for publication (in revised form) August 22, 2011; published electronically December 6, 2011. This work was partially funded by NSF grant DMS-0804696.
http://www.siam.org/journals/mms/9-4/81591.html

†Department of Applied Physics and Applied Mathematics, Columbia University, New York 10027 (gb2030@columbia.edu).
‡Département de Mathématiques et Applications, Ecole Normale Supérieure, 45 Rue d'Ulm, 75230 Paris Cedex 05, France (wjing@dms.ens.fr).

approximations of the solution without solving all the details of the microstructure [1], [10], [12]. Examples include the multiscale finite element method (MsFEM) and the finite element heterogeneous multiscale method (HMM). Such schemes are shown to perform well in the homogenization regime, in the sense that they approximate the solution to the homogenized equation without explicitly calculating any macroscopic, effective medium coefficient. Homogenization theory thus serves as a benchmark that ensures that the multiscale scheme performs well in controlled environments, with the hope that it will still perform well in noncontrolled environments, for instance, when ergodicity and stationarity assumptions are not valid.

This paper aims to present another benchmark for such multiscale numerical schemes that addresses the limiting stochasticity of the solutions. We calculate the limiting (probability) distribution of the random corrector given by the multiscale algorithm when the correlation length of the medium tends to 0 at a fixed value of the discretization size h. We then compare this distribution to the distribution of the corrector of the continuous equation. When these distributions are close, in the sense that the h-dependent distribution converges to the continuous distribution as $h \to 0$, we deduce that the multiscale algorithm asymptotically correctly captures the randomness in the solution and passes the random corrector test.

The above proposal requires a controlled environment in which the theory of correctors is available. There are very few equations for which this is the case [2], [3], [14]. In this paper, we initiate such an analysis in the simple case of the following one-dimensional second-order elliptic equation:

$$
\begin{cases} -\frac{d}{dx} a\Big(\frac{x}{\varepsilon}, \omega\Big) \frac{d}{dx} u_\varepsilon(x,\omega) = f(x), \quad x \in (0,1), \\ u_\varepsilon(0,\omega) = u_\varepsilon(1,\omega) = 0. \end{cases} \tag{1.1}
$$

Here, the diffusion coefficient $a(\frac{x}{\varepsilon},\omega)$ is modeled as a random process, where ω denotes the realization in an abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in which the random process and all limits considered here are constructed. The correlation length ε is much smaller than $L = 1$, the length of the domain, which makes the random coefficient highly oscillatory.

We chose this equation for two reasons. First, many multiscale numerical schemes have been developed to solve its generalization in higher dimensions. Second, both the homogenization and the corrector theories for this elliptic equation are well known. When $a(x,\omega)$ is stationary, ergodic, and uniformly elliptic, i.e., $\lambda \le a(x,\omega) \le \Lambda$ for almost every x and ω, then the solution u_ε converges to the following homogenized equation with deterministic and constant coefficient (this generalizes to higher dimensions as well [20], [22]):

$$
\begin{cases} -\frac{d}{dx} a^* \frac{d}{dx} u_0(x) = f(x), \quad x \in (0,1), \\ u_0(0) = u_0(1) = 0. \end{cases} \tag{1.2}
$$

The coefficient a^* is the harmonic mean of $a(x,\omega)$. Furthermore, if the deviation of $1/a(x,\omega)$ from its mean $1/a^*$ is strongly mixing with an integrable mixing coefficient, a notion that will be defined in the next section and which in particular implies that the random coefficient has short-range correlation (SRC), then the corrector theory in [6] shows that

$$\frac{u_\varepsilon - u_0}{\sqrt{\varepsilon}}(x) \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}(x; W), \tag{1.3}$$

where $\mathcal{U}(x; W)$ is a Gaussian random process that may be conveniently written as a Wiener integral; see (2.13) below. The convergence above is in the sense of distribution in the space of continuous paths, which will be denoted by $\mathcal{C}([0, 1])$. When the deviation $1/a - 1/a^*$ does not decorrelate fast enough, the normalization factor $\sqrt{\varepsilon}$ is no longer correct. In fact, for a large class of random processes with long-range correlation (LRC) defined more precisely later, it is proved in [3] that a similar convergence result holds with the following modifications: The normalization factor should be replaced by $\varepsilon^{\frac{\alpha}{2}}$, where $0 < \alpha < 1$ describes the rate of decorrelation. The limiting process is then $\mathcal{U}_H(x; W^H)$, which is defined in (2.14) and has the form of a stochastic integral with respect to fractional Brownian motion.

Let us assume that we want to solve (1.1) numerically. We denote by h the discretization size and by $u^h_\varepsilon(x)$ the solution of the scheme. The standard finite element solution of the deterministic limit (1.2) is denoted by $u^h_0(x)$. Our goal is to characterize the limiting distribution of $u^h_\varepsilon - u^h_0$ as a random process after proper rescaling by $\varepsilon^{\frac{\alpha\wedge 1}{2}}$. Here and below, we use $c \wedge d$ to denote $\min\{c, d\}$. Obviously, this normalization parameter is chosen to be consistent with the aforementioned corrector theories.

We say that a numerical procedure is consistent with the corrector theory and that it passes the corrector test when the following diagram commutes:

$$\begin{array}{ccc} \dfrac{u^h_\varepsilon - u^h_0}{\varepsilon^{\frac{\alpha\wedge 1}{2}}}(x,\omega) & \xrightarrow[\text{(i)}]{h\to 0} & \dfrac{u_\varepsilon - u_0}{\varepsilon^{\frac{\alpha\wedge 1}{2}}}(x,\omega) \\ \varepsilon \to 0 \downarrow \text{(ii)} & & \text{(iii)} \downarrow \varepsilon \to 0 \\ \mathcal{U}^h_{\alpha\wedge 1}(x; W^{\alpha\wedge 1}) & \xrightarrow[\text{(iv)}]{h\to 0} & \mathcal{U}_{\alpha\wedge 1}(x; W^{\alpha\wedge 1}). \end{array} \tag{1.4}$$

Here, $W^1 = W$ is the standard Brownian motion, whereas $W^\alpha = W^H$ is the fractional Brownian motion with Hurst index $H = 1 - \frac{\alpha}{2}$. Similarly, $\mathcal{U}_{\alpha\wedge 1}$ is defined so that $\mathcal{U}_1 = \mathcal{U}(x; W)$ and $\mathcal{U}_\alpha = \mathcal{U}_H(x; W^H)$. We use this notation to include both random processes with SRC and those with LRC. In the above diagram, there are four convergence paths to be understood in the sense of distribution in $\mathcal{C}([0, 1])$. The convergence path (iii) is the corrector theories in [6], [3]. In (i), h is sent to zero, while ε is fixed. To check (i) is a numerical analysis question without multiscale issues since the ε-scale details are resolved. Convergence in (i) can be obtained pathwise and not only in distribution (path (iv) may also be considered pathwise). The main new mathematical difficulties we address in this paper, therefore, lie in analyzing the paths (ii) and (iv).

Our main results are stated for the MsFEM [16], [17], [15], and the finite element HMM [9], [10]. The main idea of MsFEM is to use multiscale basis functions constructed from the local solutions of the elliptic operator in (1.1) on domains of size $h \ll 1$. The main idea of HMM is to construct these basis functions on possibly smaller patches. HMM thus involves heterogeneous computations on domains of size $\delta \le h$.

The analysis of MsFEM in random settings is done in [7], [11]. HMM in random media was considered in [10]. These works show that both MsFEM and HMM pass the test of homogenization theory in the sense that as $\varepsilon \to 0$, the limiting solution is a consistent discretization (i.e., with error converging to 0 as $h \to 0$) of the homogenized limit. This paper analyzes the corrector theory of both multiscale methods.

For MsFEM and HMM with the parameter choice $\delta = h$ (these two methods are then the same in dimension $n = 1$), the above diagram commutes. Path (i) holds as a result of standard finite element analysis. Path (ii) holds due to the self-averaging effect of integrals with oscillatory integrand $a^{-1}(x/\varepsilon) - a^{*-1}$, and $\mathcal{U}^h_{\alpha\wedge 1}$ is explicitly characterized as a Gaussian process. The expression for this process is complicated but converges to the right corrector as $h \to 0$.

For HMM with $\delta < h$, path (i) does not hold because there is no homogenization effect when $h \ll \varepsilon$. Passing the corrector test then means that (ii) followed by (iv) yields the same result as (iii). Path (ii) is obtained with $\mathcal{U}^h_{\alpha\wedge 1}$, a Gaussian process. HMM with $\delta < h$ then performs differently for SRC and LRC. In the first case, the limit process as $h \to 0$ in (iv) is $(h/\delta)^{\frac{1}{2}}\mathcal{U}$: The limit is an amplified version of the theoretical corrector when $\delta < h$. In that sense, HMM with $\delta < h$ does not pass the corrector test. In the LRC case, however, and somewhat surprisingly at first, HMM passes the corrector test for all $0 < \delta \le h$. As a consequence, HMM with $\delta < h$ finds the corrector up to a rescaling coefficient that depends on the structure of the random medium and whose estimation may prove difficult for environments in which the corrector theory is not valid or not known.

We consider a modification of the MSFEM and HMM schemes that is computationally more expensive than HMM with $\delta \ll h$ but passes the corrector test independently of the underlying random structure. The proposed corrector test and our analysis, in spite of the limited scope of the one-dimensional equation, thus provide some guidance on the accuracy one may expect from a multiscale algorithm based on a given computational cost and a given level of prior information about the heterogeneous medium (such as, e.g., its correlation properties) one is willing to make.

The rest of the paper is organized as follows. In the next section, we present our models for the random medium and formulate the main results of the paper. The derivations rely on establishing explicit expressions for $u^h_\varepsilon - u^h_0$. In section 3, we derive such expressions for a general numerical scheme satisfying specific assumptions on the structure of the associated stiffness matrix. We show how to prove (ii) using these formulas. In section 4, we show how to prove (iv) in the general setting. Both convergences in (ii) and (iv) can be viewed as weak convergences of measures in $\mathcal{C}$, which can be shown as usual by obtaining the convergence of finite-dimensional distributions and establishing tightness. In section 5, we apply the general framework to MsFEM and prove that diagram (1.4) commutes in this case. In section 6, we consider HMM in both media with SRC or LRC. We emphasize where the amplification effect comes from in the case of SRC and why this effect disappears in the case of long-range media. In section 7, we discuss methods to eliminate the amplification effect of HMM. Some conclusions are offered in section 8, while technical lemmas on fourth-order moments of random processes are postponed to the appendix.

2. Main results on the corrector test.

In this section, we introduce our hypotheses on the random media, describe the multiscale algorithms MsFEM and HMM, and formulate our main convergence results.

To simplify notation, we drop the dependency in ω when this does not cause confusion. We define $a_\varepsilon(x) = a(x/\varepsilon)$. For a function g in $L^p(D)$, we denote its norm by $\|g\|_{p,D}$. When D is the unit interval, we drop the symbol D. The natural space for (1.1) and (1.2) is the Hilbert space H^1_0. By the Poincaré inequality, the seminorm of H^1_0 defined by $|u|_{H^1,D} = \|du/dx\|_{2,D}$ is equivalent to the standard norm. We use the notation C to denote constants that may vary from line to line. When C depends only on the elliptic constants (λ, Λ), we refer to it as a *universal* constant. Finally, the Einstein summation convention is frequently used: Two repeated indices, such as in $c_i d^i$, are summed over their (natural) domain of definition.

2.1. Random media models. We model $a(x, \omega)$ as a random process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with parameter $x \in \mathbb{R}$. The coefficient $a(\frac{x}{\varepsilon}, \omega)$ in (1.1) is obtained by rescaling the spatial variable.

A process $a(x)$ is called *stationary* if the joint distribution of any finite collection of random variables $\{a(x_i, \omega)\}_{1 \le i \le n}$ is translation invariant. Let $\mathbb{E}$ denote the mathematical expectation with respect to $\mathbb{P}$. We define

$$\frac{1}{a^*} := \mathbb{E}\left\{\frac{1}{a(0,\omega)}\right\}, \qquad q(x,\omega) = \frac{1}{a(x,\omega)} - \frac{1}{a^*}. \tag{2.1}$$

Since a is stationary, a^* is a constant and $q(x, \omega)$ is a stationary process with mean zero.

The (auto)correlation function $R(x)$ of q is given by

$$R(x) := \mathbb{E}\{q(y)q(y+x)\} = \mathbb{E}\{q(0)q(x)\}. \tag{2.2}$$

We find that R is symmetric, i.e., $R(x) = R(-x)$, and is a function of positive type in the sense of [25, section XI]. By Bochner's theorem, the Fourier transform of R is nonnegative. Let us define

$$\sigma^2 = \int_{\mathbb{R}} R(x)\mathrm{d}x. \tag{2.3}$$

Then σ^2 is in $[0, \infty]$, and we assume here that $\sigma > 0$. We say that $q(x, \omega)$ has SRC if R is integrable. In this case, σ is a positive finite number. We say that q has LRC if the correlation function is not integrable.

The decay of $R(x)$ at infinity describes the two-point decorrelation rate of the process q. We also need the notion of *mixing*. Consider a Borel set $A \subset \mathbb{R}$. Let $\mathcal{F}_A$ denote the Borel sub-σ-algebra of $\mathcal{F}$ induced by $\{q(x); x \in A\}$ and L^2_A the space of $\mathcal{F}_A$-measurable and square integrable random variables. A random process $a(x, \omega)$ is said to be *ρ-mixing* if there exists some nonnegative function $\rho(r) : \mathbb{R}_+ \to \mathbb{R}_+$ with $\rho(r) \to 0$ at infinity, such that for any Borel sets A and B, the following holds:

$$\sup_{\xi \in L^2_A, \eta \in L^2_B} \frac{|\mathbb{E}\{\xi\eta\} - \mathbb{E}\{\xi\}\mathbb{E}\{\eta\}|}{\mathrm{Var}\{\xi\}^{\frac{1}{2}}\mathrm{Var}\{\eta\}^{\frac{1}{2}}} \le \rho(d(A, B)). \tag{2.4}$$

Here $d(A, B)$ is the distance between the two sets. The function $\rho(r)$ is called the ρ-mixing coefficient, and it reflects, roughly speaking, how fast random processes restricted on A and B become independent. More details on the notion of mixing can be found in [8].

Our first set of assumptions on $a(x)$ and $q(x)$ is the following. We note that the third assumption below implies the second.

(S1) The random process $a(x)$ is stationary and uniformly elliptic with constants (λ, Λ). Clearly, a^* is uniformly elliptic with the same constants.

(S2) The random process $q(x)$ has SRC, and we assume $\sigma > 0$ in (2.3).

(S3) The mixing coefficient $\rho(r)$ of $q(x)$ is decreasing in r, and $\rho^{\frac{1}{2}}(r) \in L^1(\mathbb{R}_+)$.

The above assumptions are quite general. In particular, (S3) implies ergodicity of $q(x)$, and (S1) plus ergodicity is the standard assumption for homogenization theory;

(S3) is the standard assumption to obtain the central limit theorem of oscillatory integrals with integrand $q_\varepsilon(x)$ as in (3.29).

For random media with LRC, there is no general central limit theorem. A special family of random media investigated in [3] is based on the following assumptions:

(L1) The process $q(x)$ is constructed as

$$q(x,\omega) = \Phi(g(x,\omega)), \tag{2.5}$$

where g_x is a stationary Gaussian random process with mean zero and variance one. Further, the correlation function R_g of g_x has the following asymptotic behavior:

$$R_g(\tau) \sim \kappa_g \tau^{-\alpha} \quad \text{as} \quad \tau \to \infty, \tag{2.6}$$

where $\kappa_g > 0$ is a constant and $\alpha \in (0,1)$.

(L2) The function $\Phi(x)$ satisfies $|\Phi| \le q_0$ for some constant q_0, so that the process $a(x,\varepsilon)$, constructed by the relation (2.1) for some positive constant a^*, satisfies uniform ellipticity with constants (λ, Λ).

(L3) The function Φ integrates to zero against the standard Gaussian measure:

$$\int_{\mathbb{R}} \Phi(s) e^{-\frac{s^2}{2}} ds = 0. \tag{2.7}$$

The process $q(x)$ is stationary and mean-zero. More importantly, its correlation function $R(x)$ has a similar asymptotic behavior to that in (2.6) with κ_g replaced by κ, where

$$\kappa := \kappa_g \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} s\Phi(s) e^{-\frac{s^2}{2}} ds \right)^2. \tag{2.8}$$

The integral above is assumed to be nonzero. Consequently, $R(x)$ is no longer integrable and $q(x)$ has LRC. We note that when $\alpha > 1$, the process constructed above has SRC and provides an example satisfying (S2).

2.2. The multiscale algorithms MsFEM and HMM. We briefly introduce MsFEM and HMM and leave some details to later sections. Assume $f \in L^2 \subset H^{-1}$. The weak solution to (1.1) is the unique function $u_\varepsilon \in H_0^1(0,1)$ such that

$$A_\varepsilon(u,v) = F(v) \qquad \forall\ v \in H_0^1(0,1). \tag{2.9}$$

The associated bilinear and linear forms are defined as

$$A_\varepsilon(u,v) := \int_0^1 a_\varepsilon(x) \frac{du}{dx} \cdot \frac{dv}{dx} dx, \qquad F(v) := \int_0^1 f v dx. \tag{2.10}$$

Existence and uniqueness of u_ε are guaranteed by the uniform ellipticity of $a_\varepsilon(x)$. MsFEM and HMM are finite element methods, and so the ideas are to approximate H_0^1 by finite-dimensional space and, when necessary, to approximate A_ε by an auxiliary bilinear form.

We partition the unit interval into N small intervals of size $h = 1/N$. We use this partition for MsFEM, HMM, and the standard FEM. The FEM basis functions are piecewise linear "hat functions," each of them peaking at one nodal point and vanishing at all other nodal points. Denote these hat functions by $\{\phi_0^j\}$ and denote the subspace of H_0^1 they span by V_0^h. The standard FEM approximates H_0^1 by V_0^h. The idea of MsFEM is to replace the hat basis functions by multiscale basis functions $\{\phi_\varepsilon^j\}$. For instance, ϕ_ε^j is constructed by solving the elliptic equation (1.1) locally on the support of ϕ_0^j; see (5.1) below. Denote the span of these multiscale basis functions by V_ε^h. Then, the MsFEM solution, denoted by u_ε^h, is the unique function in V_ε^h so that (2.9) holds for all $v \in V_\varepsilon^h$. The well-posedness of this weak formulation is a consequence of the uniform ellipticity of a_ε.

HMM aims to reduce the cost of MsFEM by computing heterogeneous solutions on smaller domains. Solutions and test functions are still sought in V_0^h. The bilinear form A_ε is then approximated by

$$A_\varepsilon^{h,\delta}(u,v) \coloneqq \sum_{k=1}^N \left(\frac{1}{\delta} \int_{I_k^\delta} a_\varepsilon(x) \frac{d(\mathcal{L}u)}{dx} \cdot \frac{d(\mathcal{L}v)}{dx}\, dx \right) h,$$

where I_k^δ is a small subinterval of size $\delta < h$ of the kth interval in the above partition. The operator $\mathcal{L}$ is defined below in (6.1). This linear operator optimally uses the microscale calculation on I_k^δ in the same way that ϕ_ε^j does in MsFEM. The HMM solution is the unique function $u_\varepsilon^{h,\delta}$ in V_0^h so that (2.9) holds when H_0^1 is replaced by V_0^h and A_ε is replaced by $A_\varepsilon^{h,\delta}$. The well-posedness of this formulation is less immediate and will be proved below.

Throughout this paper, for simplicity in presentation, we assume that the microscale solvers in MsFEM and HMM, i.e., (5.1) and (6.1), are exact.

2.3. Main theorems. We state our main convergence theorems in the setting of MsFEM and HMM, although they hold for more general schemes.

The first theorem analyzes MsFEM in the setting of SRC.

THEOREM 2.1. *Let u_ε and u_0 be solutions to* (1.1) *and* (1.2)*, respectively. Let u_ε^h be the solution to* (1.1) *obtained by MsFEM, and let u_0^h be the standard finite element approximation of u_0. Then we have the following:*

(i) *Suppose that $a(x)$ satisfies* (S1) *and f is continuous on* [0,1]*. Then*

$$|u_\varepsilon^h - u_\varepsilon|_{H^1} \le \frac{h}{\lambda\pi}\|f\|_2, \qquad \|u_\varepsilon^h - u_\varepsilon\|_2 \le \frac{h^2}{\lambda\pi^2}\|f\|_2. \tag{2.11}$$

Assume further that $q(x)$ satisfies (S2)*. Then*

$$\sup_{x\in[0,1]} |\mathbb{E}(u_\varepsilon^h(x) - u_0^h(x))^2| \le C\frac{\varepsilon}{h^2}\|R\|_{1,\mathbb{R}}(1+\|f\|_2), \tag{2.12}$$

where C is a universal constant and R is the correlation function of q as defined in (2.2)*.*

(ii) *Now assume further that $q(x)$ satisfies* (S3)*. Then,*

$$\frac{u_\varepsilon^h(x) - u_0^h(x)}{\sqrt{\varepsilon}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma \int_0^1 L^h(x,t)\, dW_t =: \mathcal{U}^h(x;W). \tag{2.13}$$

The constant σ is defined in (2.3) *and W is the standard Wiener process. The function $L^h(x,t)$ is explicitly given in* (3.20)*. The convergence is in distribution in the space $\mathcal{C}$.*

(iii) *Now let h go to zero; we have*

$$\mathcal{U}^h(x;W) \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}(x;W) := \sigma \int_0^1 L(x,t)\,dW_t. \tag{2.14}$$

The Gaussian process $\mathcal{U}(x;W)$ was characterized in [6]*. The kernel $L(x,t)$ is defined as*

$$L(x,t) = \mathbf{1}_{[0,x]}(t)\Big(\int_0^1 F(s)ds - F(t)\Big) + x\Big(F(t) - \int_0^1 F(s)ds\Big)\mathbf{1}_{[0,1]}(t). \tag{2.15}$$

Here and below, $\mathbf{1}$ *is the indicator function and* $F(t) = \int_0^t f(s)\mathrm{d}s$.

Equivalently, the theorem says the diagram in (1.4) commutes when q has SRC and that MsFEM passes the corrector test in this setting.

To prove (iv) of the diagram, we recast $L(x,t)$ as

$$L(x,t) = a^* \frac{\partial}{\partial y} G_0(x,t) \cdot a^* \frac{\partial}{\partial x} u_0(t). \tag{2.16}$$

Here G_0 is the Green's function of (1.2). It has the following expression:

$$G_0(x,y) = \begin{cases} a^{*-1}x(1-y), & x \le y, \\ a^{*-1}(1-x)y, & x > y. \end{cases} \tag{2.17}$$

In particular, G_0 is Lipschitz continuous in each variable, while the other is kept fixed.

The next theorem accounts for MsFEM in media with LRC. We recall that the random process $q(x)$ below is constructed as a function of a Gaussian process.

THEOREM 2.2. *Let u_ε, u_0, u_ε^h, and u_0^h be defined as in the previous theorem. Let $q(x,\omega)$ and $a(x,\omega)$ be constructed as in* (L1)–(L3)*. Then we have the following:*

(i)

$$\sup_{x\in[0,1]} |\mathbb{E}(u_\varepsilon^h(x) - u_0^h(x))^2| \le C\frac{1}{h}\left(\frac{\varepsilon}{h}\right)^\alpha \tag{2.18}$$

for some constant C depending on (λ, Λ), κ, α, and f.

(ii) *As ε goes to zero while h is fixed, we have*

$$\frac{u_\varepsilon^h(x) - u_0^h(x)}{\varepsilon^{\frac{\alpha}{2}}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}_H^h(x;W^H) := \sigma_H \int_0^1 L^h(x,t)\,dW_t^H. \tag{2.19}$$

Here $H = 1 - \frac{\alpha}{2}$, and W_t^H is the standard fractional Brownian motion with Hurst index H. The constant σ_H is defined as $\sqrt{\kappa/H(2H-1)}$. The function $L^h(x,t)$ is defined as in the previous theorem.

(iii) *As h goes to zero, we have*

$$\mathcal{U}_H^h(x;W^H) \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}_H(x;W^H) := \sigma_H \int_0^1 L(x,t)\,dW_t^H. \tag{2.20}$$

As before, this theorem says the diagram in (1.4) commutes in the current case. In particular, $\alpha < 1$, and the scaling is $\varepsilon^{\frac{\alpha}{2}}$. Thus MsFEM passes the corrector test for both SRC and LRC. The stochastic integrals in (2.19) and (2.20) have fractional Brownian motions as integrators. We give a short review of such integrals in Appendix A. A good reference is [23].

The next theorem addresses the convergence properties of HMM.

THEOREM 2.3. *Let* u_ε *and* u_0 *be the solutions to* (1.1) *and* (1.2), *respectively. Let* $u_\varepsilon^{h,\delta}$ *be the HMM solution and* u_0^h *the standard finite element approximation of* u_0.

(i) *Suppose that the random processes* $a(x)$ *and* $q(x)$ *satisfy* (S1)–(S3). *Then*

$$\frac{u_\varepsilon^{h,\delta}(x) - u_0^h(x)}{\sqrt{\varepsilon}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}^{h,\delta}(x; W) \xrightarrow[h\to 0]{\text{distribution}} \sqrt{\frac{h}{\delta}}\,\mathcal{U}(x; W). \tag{2.21}$$

Here, $\mathcal{U}^{h,\delta}(x; W)$ *is as in* (2.13) *with* L^h *replaced by* $L^{h,\delta}(x,t)$, *which is defined in* (6.7) *below. The process* $\mathcal{U}(x; W)$ *is as in* (2.14).

(ii) *Suppose instead that the random processes* $a(x)$ *and* $q(x)$ *satisfy* (L1)–(L3). *Then*

$$\frac{u_\varepsilon^{h,\delta}(x) - u_0^h(x)}{\varepsilon^{\frac{\alpha}{2}}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \mathcal{U}_H^{h,\delta}(x; W^H) \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}_H(x; W^H). \tag{2.22}$$

Here, $\mathcal{U}_H^{h,\delta}(x; W^H)$ *is as in* (2.19) *with* L^h *replaced by* $L^{h,\delta}$, *and* $\mathcal{U}_H(x; W^H)$ *is as in* (2.20).

HMM is computationally less expensive than MsFEM when δ is much smaller than h. However, the theorem implies that this advantage comes at a price: When the random process $q(x)$ has SRC, HMM with $\delta < h$ amplifies the variance of the corrector. We will discuss methods to eliminate this effect in section 7. In the case of LRC, however, HMM does pass the corrector test even when $\delta \ll h$.

Intuitively, averaging occurs at the small-scale $\delta \ll h$ for SRC. Since HMM performs calculations on a small fraction of each interval h, each integral needs to be rescaled by h/δ to capture the correct mean, which overamplifies the size of the fluctuations. In the case of LRCs, the self-similar structure of the limiting process shows that the convergence to the Gaussian process occurs simultaneously at all scales (larger than ε) and hence at the macroscopic scale. HMM may then be seen as a collocation method (with grid size h), which does capture the main features of the random integrals.

The amplification of the random fluctuations might be rescaled to provide the correct answer. The main difficulty is that the rescaling factor depends on the structure of the random medium and thus requires prior information or additional estimations about the medium. For general random media with no clear scale separation or no stationarity assumptions, the definition of such a rescaling coefficient might be difficult. In section 7, we present a hybrid method between HMM and MsFEM that is computationally less expensive than the MsFEM presented above, while still passing the corrector test.

3. Expression for the corrector and convergence as $\varepsilon \to 0$.

The starting point to prove the main theorems is to derive a formula for the corrector $u_\varepsilon^h - u_0^h$. This is done by inverting the discrete systems yielded from the multiscale schemes. The goal is to write the corrector as a stochastic integral linear in the random coefficient, plus error terms that are negligible in the limit; see Proposition 3.2 below.

3.1. General finite element based multiscale schemes.

Almost all finite element based multiscale schemes for (1.1) have the same main premise: In the weak formulation (2.9), we approximate H_0^1 by a finite-dimensional space and, if necessary, also approximate the bilinear form.

To describe the choices of the finite spaces, we choose a uniform partition of the unit interval into N subintervals with size $h = 1/N$. Let x_k denote the kth grid point, with $x_0 = 0$ and $x_N = 1$, and I_k the interval (x_{k-1}, x_k). To simplify notation in the general setting, we still denote by V_ε^h the finite space and by $\{\phi_\varepsilon^j\}_{j=1}^{N-1}$ the basis functions. They do not necessarily coincide with those in MsFEM.

In a general scheme, the bilinear form in (2.9) might be modified. Nevertheless, to simplify notation, we still denote it as A_ε. The solution obtained from the scheme is then $u_\varepsilon^h \in V_\varepsilon^h$ such that

$$A_\varepsilon(u_\varepsilon^h, v) = F(v) \qquad \forall\ v \in V_\varepsilon^h. \tag{3.1}$$

Since V_ε^h is finite-dimensional, the above condition amounts to a linear system, which is obtained by putting $u_\varepsilon^h = U_j^\varepsilon \phi_\varepsilon^j$, and by requiring the above equation to hold for all basis functions. The linear system is

$$A_\varepsilon^h U^\varepsilon = F^\varepsilon. \tag{3.2}$$

Here, the vector U^ε is a vector in $\mathbb{R}^{N-1}$, and it has entries U_j^ε. The load vector F^ε is also in $\mathbb{R}^{N-1}$ and has entries $F(\phi_\varepsilon^j)$. The stiffness matrix A_ε^h is an $N-1$ by $N-1$ matrix, and its entries are $A_\varepsilon(\phi_\varepsilon^i, \phi_\varepsilon^j)$. Our main assumptions on the basis functions and the stiffness matrix are the following.

(N1) For any $1 \le j \le N-1$, the basis function ϕ_ε^j is supported on $I_j \cup I_{j+1}$, and it takes the value δ_j^i at nodal points $\{x_i\}$. Here δ_j^i is the Kronecker symbol.

(N2) The matrix A_ε^h is symmetric and tridiagonal. In addition, we assume that there exists a vector $b_\varepsilon \in \mathbb{R}^N$ with entries $\{b_\varepsilon^k\}_{k=1}^N$, so that $A_{\varepsilon ii+1}^h = -b_\varepsilon^{i+1}$ for any $i = 1, \ldots, N-2$ and

$$A_{\varepsilon ii}^h = -(A_{\varepsilon ii-1}^h + A_{\varepsilon ii+1}^h), \qquad i = 1, \ldots, N-1. \tag{3.3}$$

In other words, the ith diagonal entry of A_ε^h is the negative sum of its neighbors in each row. Here, $A_{\varepsilon 01}^h$ and $A_{\varepsilon N-1N}^h$ are not matrix elements and are set to be b_ε^1 and b_ε^N, respectively.

(N3) On each interval I_j for $j = 1, \ldots, N$, the only two basis functions that are nonzero are ϕ_ε^j and ϕ_ε^{j-1}, and they sum to one; i.e., $\phi_\varepsilon^j + \phi_\varepsilon^{j-1} = 1$. Equivalently, we have

$$\phi_\varepsilon^j(x) = \mathbf{1}_{I_j}\tilde{\phi}_\varepsilon^j(x) + \mathbf{1}_{I_{j+1}}(x)[1 - \tilde{\phi}^{j+1}(x)] \tag{3.4}$$

for some functions $\{\tilde{\phi}_\varepsilon^k(x)\}_{k=1}^N$, each of them defined only on I_j with boundary value 0 at the left end point and 1 at the right.

As we shall see for MsFEM, (N3) implies (N2) when the bilinear form is symmetric. The special tridiagonal structure of A_ε^h in (N2) simplifies the calculation of its action on a vector. Let U be any vector in $\mathbb{R}^{N-1}$; we have

$$(A_\varepsilon^h U)_i = -D^+(b_\varepsilon^i D^- U_i), \qquad i = 1, \ldots, N-1. \tag{3.5}$$

Here, the symbol D^- denotes the backward difference operator, which is defined, together with the forward difference operator D^+, as

$$(D^- U)_k = U_k - U_{k-1}, \qquad (D^+ U)_k = U_{k+1} - U_k. \tag{3.6}$$

The equality (3.5) is easy to check, and to make sense of the case when i equals 1 or N, we need to extend the definition of U by setting U_0 and U_N to zero. This formula has been used, for example, in [17]. It is a very useful tool in the forthcoming computations.

3.2. The expression of the corrector. Now we derive an expression of the corrector, i.e., the difference between u_ε^h, the solution to (1.1) obtained from the above scheme, and u_0^h, the standard FEM solution to (1.2).

The function $u_0^h(x)$ is obtained from a weak formulation similar to (2.9) with a_ε replaced by a^*, and H_0^1 replaced by V_0^h, the space spanned by hat functions $\{\phi_0^j\}$. Clearly, these basis functions satisfy (N1) and (N3). Let A_0^h denote the associated stiffness matrix; then one can verify that it satisfies (N2). In fact, the vector b is given by $b_0^k = a^*/h$. Now $u_0^h(x)$ is simply $U_j^0\phi_0^j$, where U^0 solves

$$A_0^h U^0 = F^0. \tag{3.7}$$

Subtracting this equation from (3.2), we obtain

$$A_0^h(U^\varepsilon - U^0) = (F^\varepsilon - F^0) - (A_\varepsilon^h - A_0^h)U^\varepsilon.$$

Let G_0^h denote the inverse of the matrix A_0^h. We have

$$U^\varepsilon - U^0 = G_0^h(F^\varepsilon - F^0) - G_0^h(A_\varepsilon^h - A_0^h)U^\varepsilon.$$

Since both A_ε^h and A_0^h satisfy (N2), the difference $A_\varepsilon^h - A_0^h$ acts on vectors in the same manner as in (3.5). Since both $\{\phi_\varepsilon^j\}$ and $\{\phi_0^j\}$ satisfy (N3), we verify that

$$(F^\varepsilon - F^0)_j = -D^+(\tilde{F}_j^\varepsilon - \tilde{F}_j^0), \qquad \tilde{F}_j^\varepsilon := \int_{I_j} f(t)\tilde{\phi}_\varepsilon^j(t)\,dt.$$

Using these difference forms, we have

$$\begin{aligned} U_j^\varepsilon - U_j^0 &= -\sum_{m=1}^{N-1}(G_0^h)_{jm}(D^+(\tilde{F}^\varepsilon - \tilde{F}^0)_m - D^+((b_\varepsilon^m - b_0^m)D^- U_m^\varepsilon)) \\ &= \sum_{k=1}^{N} D^- G_{0jk}^h((\tilde{F}^\varepsilon - \tilde{F}^0)_k - (b_\varepsilon^k - b_0^k)D^- U_k^\varepsilon). \end{aligned} \tag{3.8}$$

The second equality is obtained from summation by parts. Note that we have extended the definitions of U^ε and U^0 so that they equal zero when the index is 0 or N. Similarly, $(G_0^h)_{j0}$ and $(G_0^h)_{jN}$ are zero as well.

The vector $U^\varepsilon - U^0$ is the corrector evaluated at the nodal points. We have the following control of its ℓ^2 norm under some assumptions on the statistics of $\{b_\varepsilon^k\}$ and $\{\phi_\varepsilon^j\}$.

Proposition 3.1. *Let U^ε and U^0 be as above. Let the basis functions $\{\phi_\varepsilon^j\}$ and the stiffness matrix A_ε^h satisfy* (N1)–(N3). *Suppose also that*

$$\sup_{1\le k\le N} |D^- U_k^\varepsilon| \le C\|f\|_2 h^{\frac{1}{2}} \tag{3.9}$$

for some universal constant C.

(i) *Suppose further that for any $k = 1, \ldots, N$, and any $x \in I_k$, we have*

$$\mathbb{E}(\tilde{\phi}^k_\varepsilon(x) - \tilde{\phi}^k_0(x))^2 \le C\frac{\varepsilon}{h}\|R\|_{1,\mathbb{R}}, \tag{3.10}$$

and

$$\mathbb{E}(b^k_\varepsilon - b^k_0)^2 \le C\frac{\varepsilon}{h^3}\|R\|_{1,\mathbb{R}} \tag{3.11}$$

for some universal constant C. Then we have

$$\mathbb{E}\| U^\varepsilon - U^0\|^2_{\ell^2} \le C\frac{\varepsilon}{h^3}\|R\|_{1,\mathbb{R}}(1+\|f\|_2) \tag{3.12}$$

for some universal C.

(ii) *Suppose instead that the right-hand side of* (3.10) *is $C(\frac{\varepsilon}{h})^\alpha$, and the right-hand side of* (3.11) *is $C\frac{1}{h^2}(\frac{\varepsilon}{h})^\alpha$. Then the estimate in* (3.12) *should be changed to $C\frac{1}{h^2}(\frac{\varepsilon}{h})^\alpha$.*

The assumption (3.9) essentially says that u^h_ε should have a Hölder regularity. Suppose the weak formulation associated with the multiscale scheme admits a unique solution u^h_ε such that $\|u^h_\varepsilon\|_{H^1} \le C\|f\|_2$. Then by Morrey's inequality [13, p. 266], $u^h_\varepsilon \in C^{0,\frac{1}{2}}$ in one dimension. Consequently, (3.9) holds.

For MsFEM, we have a better estimate: $|D^- U^\varepsilon_k| \le Ch$ due to a superconvergence result; see (5.3). Therefore, the estimate in (3.12) can be improved to be $C\frac{\varepsilon}{h^2}$ in case (i) and $C\frac{1}{h}(\frac{\varepsilon}{h})^\alpha$ in case (ii).

Item (i) of this proposition is useful when the random medium $a(x)$, or equivalently $q(x)$, has SRC, while item (ii) is useful in the case of LRC. The constant C in the second item depends on (λ,Λ), f, and R_g but not on h.

Proof. To prove (i), we use a superconvergence result, which we prove in section 5.1, to get $|D^- G^h_{0jk}| \le Ch$. Using this estimate together with (3.9) and (3.8), we have

$$\mathbb{E}|U^\varepsilon_j - U^0_j|^2 \le Ch\sum_{k=1}^N \mathbb{E}|\tilde{F}^\varepsilon_k - \tilde{F}^0_k|^2 + Ch^2\sum_{k=1}^N \mathbb{E}|b^k_\varepsilon - b^k_0|^2.$$

For the second term, we use (3.11) and obtain

$$\sum_{k=1}^N \mathbb{E}|b^k_\varepsilon - b^k_0|^2 \le \sum_{k=1}^N C\frac{\varepsilon}{h^3}\|R\|_{1,\mathbb{R}} = C\frac{\varepsilon}{h^4}\|R\|_{1,\mathbb{R}}. \tag{3.13}$$

For the other term, an application of Cauchy–Schwarz to the definition of $\tilde{F}^\varepsilon$ yields

$$|\tilde{F}^\varepsilon_k - \tilde{F}^0_k|^2 \le \|f\|^2_{2,I_k}\|\tilde{\phi}^k_\varepsilon - \tilde{\phi}^k_0\|^2_{2,I_k}.$$

Using (3.10), we have

$$\mathbb{E}\|\tilde{\phi}^k_\varepsilon - \tilde{\phi}^k_0\|^2_{2,I_k} = \int_{I_k}\mathbb{E}(\tilde{\phi}^k_\varepsilon - \tilde{\phi}^k_0)^2(x)\,dx \le C\frac{\varepsilon}{h}\cdot h\|R\|_{1,\mathbb{R}} = C\varepsilon\|R\|_{1,\mathbb{R}}. \tag{3.14}$$

Therefore, we have

$$\mathbb{E}|U^\varepsilon_j - U^0_j|^2 \le \left(Ch\sum_{k=1}^N \|f\|_{2,I_k}\varepsilon + Ch^2\frac{\varepsilon}{h^4}\right)\|R\|_{1,\mathbb{R}} \le C\frac{\varepsilon}{h^2}\|R\|_{1,\mathbb{R}}(1+\|f\|_2).$$

Note that this estimate is uniform in j. Sum over j to complete the proof of (i).

Proof of item (ii) follows in exactly the same way, using the corresponding estimates. ∎

Now, the corrector in this general multiscale numerical scheme is

$$\begin{aligned} u_\varepsilon^h(x) - u_0^h(x) &= U_j^\varepsilon \phi_\varepsilon^j(x) - U_j^0 \phi_0^j(x) \\ &= (U^\varepsilon - U^0)_j \phi_0^j(x) + U_j^0(\phi_\varepsilon^j - \phi_0^j)(x) + (U^\varepsilon - U^0)_j(\phi_\varepsilon^j - \phi_0^j)(x). \end{aligned} \tag{3.15}$$

We call the three terms on the right-hand side $K_i(x)$, $i = 1, 2, 3$. Now $K_1(x)$ is the piecewise interpolation of the corrector evaluated at the nodal points, $K_2(x)$ is the corrector due to different choices of basis functions, and $K_3(x)$ is much smaller due to the previous proposition and (3.10). Our analysis shows that $K_1(x)$ and $K_2(x)$ contribute to the limit when $\varepsilon \to 0$ while h is fixed, but only a part of $K_1(x)$ contributes to the limit when $h \to 0$.

Due to self-averaging effect, which is made precise in Lemma A.1, integrals of $q_\varepsilon(x)$ are small. Therefore, our goal is to decompose the above expression into two terms: a leading term that is an oscillatory integral against q_ε, and a remainder term that contains multiple oscillatory integrals.

Proposition 3.2. *Assume that u_ε^h is the solution to* (1.1) *obtained from a multiscale scheme, which satisfies* (N1)–(N3) *and has basis functions $\{\phi_\varepsilon^j\}$, and that u_0^h is the solution of* (1.2) *obtained by the standard FEM with hat basis functions $\{\phi_0^j\}$. Let b_ε and b_0 denote the vectors in* (N2) *of these methods. Suppose that* (3.9) *holds and that for any $k = 1, \dots, N$, we have*

$$\begin{aligned} \tilde{\phi}_\varepsilon^k(t) - \tilde{\phi}_0^k(t) &= [1 + \tilde{r}_{1k}] \frac{a^*}{h} \left(\int_{x_{k-1}}^t q_\varepsilon(s)\,ds - \frac{t - x_{k-1}}{h} \int_{x_{k-1}}^{x_k} q_\varepsilon(s)\,ds \right), \\ b_\varepsilon^k - b_0^k &= [1 + \tilde{r}_{2k}] \left(-\frac{a^{*2}}{h^2} \int_{x_{k-1}}^{x_k} q_\varepsilon(t)\,dt \right) \end{aligned} \tag{3.16}$$

for some random variables $\tilde{r}_{1k}$ and $\tilde{r}_{2k}$.

(i) *Assume that $q(x)$ has SRC, i.e., satisfies* (S2), *and that*

$$\sup_{1 \le k \le N} \max\{\mathbb{E}|\tilde{r}_{1k}|^2, \mathbb{E}|\tilde{r}_{2k}|^2\} \le C \frac{\varepsilon}{h} \|R\|_{1,\mathbb{R}} \tag{3.17}$$

for some universal constant C. Then, the corrector can be written as

$$u_\varepsilon^h(x) - u_0^h(x) = \int_0^1 L^h(x,t) q_\varepsilon(t)\,dt + r_\varepsilon^h(x). \tag{3.18}$$

Furthermore, the remainder $r_\varepsilon^h(x)$ satisfies

$$\sup_{x \in [0,1]} \mathbb{E}|r_\varepsilon^h(x)| \le C \frac{\varepsilon}{h^{5/2}} \|R\|_{1,\mathbb{R}} (1 + \|f\|_2) \tag{3.19}$$

for some universal constant C. The function $L^h(x,t)$ is the sum of L_1^h and L_2^h defined by

$$L_1^h(x,t) = \sum_{k=1}^{N} \mathbf{1}_{I_k}(t) \frac{a^* D^- G_0^h(x,x_k)}{h}$$
$$\times \left(\frac{a^* D^- U_k^0}{h} + \int_t^{x_k} f(s)ds - \int_{x_{k-1}}^{x_k} f(s)\tilde{\phi}_0^k(s)ds \right),$$

$$L_2^h(x,t) = \frac{a^*}{h} D^- U_{0j(x)}^h \left(\mathbf{1}_{[x_{j(x)-1},x]}(t) - \frac{x - x_{j(x)-1}}{h} \mathbf{1}_{[x_{j(x)-1},x_{j(x)}]}(t) \right). \tag{3.20}$$

Given x, the index $j(x)$ is the unique one so that $x_{j(x)-1} < x \le x_{j(x)}$. The function $G_0^h(x,x_k)$ is defined as

$$G_0^h(x,x_k) = \sum_{j=1}^{N-1} G_{0jk}^h \phi_0^j(x). \tag{3.21}$$

G_0^h is the interpolation in V_0^h using the discrete Green's function of standard FEM.

(ii) *Assume that $q(x)$ has LRC, i.e.,* (L1)–(L3) *are satisfied, and that the estimate in* (3.17) *is $C(\frac{\varepsilon}{h})^\alpha$. Then the same decomposition holds; the expression of $L^h(x,t)$ remains the same, but the estimate in* (3.19) *should be replaced by $C\frac{1}{h^{3/2}}(\frac{\varepsilon}{h})^\alpha$.*

Due to the superconvergent result in section 5.1, the function $G_0^h(x,x_k)$ above is exactly the Green's function evaluated at (x,x_k). This can be seen from the facts that they agree at nodal points and are both piecewise linear and continuous.

Proof. We only present the proof of item (i). Item (ii) follows in exactly the same way. We point out that the assumption (3.16) and the estimates (3.17) imply (3.10) and (3.11) thanks to Lemma A.1.

The idea is to extract the terms in the expression (3.15) that are linear in q_ε. For $K_1(x)$, we use (3.8) and write

$$K_1(x) \approx \sum_{j=1}^{N-1} \sum_{k=1}^{N} D^- G_{0jk}^h (\tilde{F}_k^\varepsilon - \tilde{F}_k^0 - (b_\varepsilon^k - b_0^k) D^- U_k^0) \phi_0^j(x)$$
$$= \sum_{k=1}^{N} D^- G_0^h(x,x_k)(\tilde{F}_k^\varepsilon - \tilde{F}_k^0 - (b_\varepsilon^k - b_0^k) D^- U_k^0).$$

Note that the expression above is an approximation because we have changed $D^- U^\varepsilon$ on the right-hand side of (3.8) to $D^- U^0$. The error is

$$r_{11}^h(x) = -\sum_{k=1}^{N} D^- G_0^h(x,x_k)(b_\varepsilon^k - b_0^k) D^-(U^\varepsilon - U^0)_k. \tag{3.22}$$

Estimating $|D^- G_0^h|$ by Ch and using Cauchy–Schwarz on the sum over k and (3.13) and (3.12), we verify that $\mathbb{E}|r_{11}^h(x)| \le C\varepsilon h^{-5/2} \|R\|_{1,\mathbb{R}}(1 + \|f\|_2)$.

Using the expressions of $\tilde{\phi}_\varepsilon$ and b_ε, and the estimates of the higher order terms in them, (3.16), we can further approximate $K_1(x)$ by

$$K_1(x) \approx \sum_{k=1}^{N} D^- G_0^h(x, x_k)\Big(\int_{x_{k-1}}^{x_k} f(t)\frac{a^*}{h}\Big[\int_{x_{k-1}}^{t} q_\varepsilon(s)ds - \frac{t-x_{k-1}}{h}\int_{x_{k-1}}^{x_k} q_\varepsilon(s)ds\Big]dt + \frac{a^{*2}}{h^2}D^- U_k^0 \int_{x_{k-1}}^{x_k} q_\varepsilon(t)dt\Big).$$

The error in this approximation is

$$r_{12}^h(x) = \sum_{k=1}^{N} D^- G_0^h(x, x_k)\Big(\tilde{r}_{1k}\int_{x_{k-1}}^{x_k} f(t)\frac{a^*}{h}\Big[\int_{x_{k-1}}^{t} q_\varepsilon(s) - \frac{t-x_{k-1}}{h}\int_{x_{k-1}}^{x_k} q_\varepsilon(s)\Big]dt + \tilde{r}_{2k}\frac{a^{*2}}{h^2}D^- U_k^0 \int_{x_{k-1}}^{x_k} q_\varepsilon(t)dt\Big). \tag{3.23}$$

Using Lemma A.1, (3.17), and Cauchy–Schwarz, we have

$$\mathbb{E}\Big|\tilde{r}_{1k}\int_{I_k} q_\varepsilon(s)ds\Big| \le C\varepsilon.$$

Using this estimate, we verify that the mean of the absolute value of the first term in r_{12}^h is bounded by $C\varepsilon\|f\|_2\|R\|_{1,\mathbb{R}}$. A similar estimate with $|D^- U_k^0| \le Ch$ (due to superconvergence) shows that the second term in r_{12}^h has absolute mean bounded by $C\varepsilon h^{-1}\|R\|_{1,\mathbb{R}}$. Therefore, we have $\mathbb{E}|r_{12}^h(x)| \le C\varepsilon h^{-1}(1+\|f\|_2)\|R\|_{1,\mathbb{R}}$. We remark also that in the case of LRCs, we should apply Lemma A.2 instead.

Moving on to $K_2(x)$, we observe that for fixed x, $K_2(x)$ reduces to a sum over at most two terms, due to the fact that ϕ_ε^j and ϕ_0^j have local support only. Let $j(x)$ be the index so that $x \in (x_{j(x)-1}, x_{j(x)}]$. We have

$$\begin{aligned} K_2(x) &= \sum_{j=1}^{N} D^- U_j^0(\tilde{\phi}_\varepsilon^j - \tilde{\phi}_0^j)(x) = D^- U_{j(x)}^0(\tilde{\phi}_\varepsilon^{j(x)} - \tilde{\phi}_0^{j(x)})(x) \\ &\approx D^- U_{j(x)}^0 \frac{a^*}{h}\Big(\int_{x_{j(x)-1}}^{x} q_\varepsilon(t)dt - \frac{x - x_{j(x)-1}}{h}\int_{x_{j(x)-1}}^{x_{j(x)}} q_\varepsilon(t)dt\Big). \end{aligned}$$

In the second step above, we used the decomposition of $\tilde{\phi}_\varepsilon$ again. The error we make in this step is

$$r_2^h(x) = \tilde{r}_{1j(x)}\frac{a^* D^- U_{j(x)}^0}{h}\Big(\int_{x_{j(x)-1}}^{x} q_\varepsilon(t)dt - \frac{x - x_{j(x)-1}}{h}\int_{x_{j(x)-1}}^{x_{j(x)}} q_\varepsilon(t)dt\Big). \tag{3.24}$$

We verify again that $\mathbb{E}|r_2^h(x)| \le C\varepsilon\|R\|_{1,\mathbb{R}}$.

Now for $K_3(x)$, we use Cauchy–Schwarz and have

$$\mathbb{E}|K_3(x)| \le \mathbb{E}\Big(\sum_{j=1}^{N-1}|U_{\varepsilon j}^h - U_{0j}^h|^2\Big)^{\frac{1}{2}}\Big(\sum_{j=1}^{N-1}(\phi_\varepsilon^j - \phi_0^j)^2(x)\Big)^{\frac{1}{2}} \le C\frac{\varepsilon}{h^2}\|R\|_{1,\mathbb{R}}(1+\|f\|_2). \tag{3.25}$$

The last inequality is due to (3.12) and (3.10).

In the approximations of $K_1(x)$ and $K_2(x)$, we change the order of summation and integration. We find that $K_1(x)$ is then $\int_0^1 L_1^h(x,t)q_\varepsilon(t)dt$ plus the error term $r_{11}^h + r_{12}^h$, and $K_2(x)$ is $\int_0^1 L_2^h(x,t)q_\varepsilon(t)dt$ plus the error term r_2^h. Therefore, we proved (3.18) with $r_\varepsilon^h(x) = r_{11}^h + r_{12}^h + r_2^h + K_3(x)$. The estimates above for these error terms are uniform in x, verifying (3.19). ∎

3.3. Weak convergence of the corrector of a multiscale scheme. In this section, we characterize the limit of the corrector $u_\varepsilon^h - u_0^h$, with proper scaling, in the multiscale scheme when ε is sent to zero. As we have seen, the scaling depends on the correlation range of the random media.

Proposition 3.3. *Let u_ε^h be the solution to* (1.1) *given by a multiscale scheme that satisfies* (N1)–(N3)*. Suppose* (3.9) *holds. Let u_0^h be the standard FEM solution to* (1.2)*.*

(i) *Suppose that $q(x)$ satisfies* (S1)–(S3) *and that the conditions of item* (i) *in Proposition* 3.2 *hold. Then,*

$$\frac{u_\varepsilon^h - u_0^h}{\sqrt{\varepsilon}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma \int_0^1 L^h(x,t)\,dW_t. \tag{3.26}$$

(ii) *Suppose that $q(x)$ satisfies* (L1)–(L3) *and that the conditions of item* (ii) *in Proposition* 3.2 *hold. Then,*

$$\frac{u_\varepsilon^h - u_0^h}{\varepsilon^{\frac{\alpha}{2}}} \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma_H \int_0^1 L^h(x,t)\,dW_t^H. \tag{3.27}$$

The real number σ is defined in (2.3) *and σ_H is defined in Theorem* 2.2*.*

These results allow us to prove the weak convergence in step (ii) of the diagram in (1.4) for fairly general schemes. A standard method to attain such weak convergence results is to use the following proposition; see [19, p. 64].

Proposition 3.4. *Suppose $\{M_\varepsilon\}$ with $\varepsilon \in (0,1)$ is a family of random processes with values in the space of continuous functions $\mathcal{C}$ and $M_\varepsilon(0) = 0$. Then M_ε converges in distribution to M_0 as $\varepsilon \to 0$ if the following hold:*

(i) *(Finite-dimensional distributions) For any $0 \le x_1 \le \cdots \le x_k \le 1$, the joint distribution of $(M_\varepsilon(x_1), \ldots, M_\varepsilon(x_k))$ converges to that of $(M_0(x_1), \ldots, M_0(x_k))$ as $\varepsilon \to 0$.*

(ii) *(Tightness) The family $\{M_\varepsilon\}_{\varepsilon\in(0,1)}$ is a tight sequence of random processes in $\mathcal{C}(I)$. A sufficient condition is the Kolmogorov criterion: $\exists \delta, \beta, C > 0$ such that*

$$\mathbb{E}\{|M_\varepsilon(s) - M_\varepsilon(t)|^\beta\} \le C|t-s|^{1+\delta}, \tag{3.28}$$

uniformly in ε and $t, s \in (0,1)$.

The standard Kolmogorov criterion for tightness requires the existence of $t \in [0,1]$ and some exponent ν so that $\sup_\varepsilon \mathbb{E}|M_\varepsilon(t)|^\nu \le C$ for C independent of ε and ν. In our cases, since $M_\varepsilon(0) = 0$ for all ε, this condition is always satisfied.

We will prove item (i) of Proposition 3.3 in detail; proof of item (ii) follows in the same way, so we only point out the necessary modifications. Recall the decomposition in (3.18). Let I_ε denote the first member on the right-hand side of this equation, i.e., the oscillatory integral. Let $\mathcal{I}^h$ denote the right-hand side of (3.26). The strategy in the case of SRC is to show that $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon\}$ converges in distribution in $\mathcal{C}$ to the target process $\mathcal{I}^h$, while $\{\varepsilon^{-\frac{1}{2}} r_\varepsilon^h\}$ converges in distribution in $\mathcal{C}$ to the zero function. Since the zero process is deterministic, the convergence in fact holds in probability; see [5, p. 27]. Then (3.26) follows.

Proof. Convergence of $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon\}$. We first check that the finite-dimensional distributions of $I_\varepsilon(x)$ converge to those of $\mathcal{I}^h(x)$. Using characteristic functions, this amounts to showing

$$\mathbb{E}\exp\left(\frac{i}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)\sum_{j=1}^n \xi_j L^h(x^j,t)\,dt\right)\xrightarrow{\varepsilon\to 0}\mathbb{E}\exp\left(i\sigma\int_0^1\sum_{j=1}^n \xi_j L^h(x^j,t)\,dW_t\right)$$

for any positive integer n, and any n-tuples $(x^1,\dots,x^n)$ and $(\xi_1,\dots,\xi_n)$. We set $m(t)=\sum_{j=1}^n \xi_j L^h(x^j,t)$. The convergence above is proved if we can show

$$\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)m(t)\,dt \xrightarrow[\varepsilon\to 0]{\text{distribution}} \sigma\int_0^1 m(t)\,dW_t \tag{3.29}$$

for any $m(t)$ that is square integrable on $[0,1]$. Indeed, this convergence holds as long as $q(x,\omega)$ is a stationary mean-zero process that admits an integrable ρ-mixing coefficient $\rho(r)\in L^1(\mathbb{R})$. This result is more or less standard, and a proof can be found in [2, Theorem 3.7 and its proof]. Our assumptions (S1)–(S3) guarantee the existence of such a $\rho(r)$. Therefore, we proved the convergence of the finite distributions of $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon\}$.

Next, we establish tightness of $\{\varepsilon^{-\frac{1}{2}} I_\varepsilon(x)\}$ by verifying (3.28). Consider the fourth moments and recall $L^h = L_1^h + L_2^h$ in (3.20); we have

$$\begin{aligned}\mathbb{E}(I_\varepsilon(x)-I_\varepsilon(y))^4 \le 8\Bigg\{&\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)(L_1^h(x,t)-L_1^h(y,t))\,dt\right)^4\\ &+\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)(L_2^h(x,t)-L_2^h(y,t))\,dt\right)^4\Bigg\}.\end{aligned}\tag{3.30}$$

We estimate the two terms on the right separately. For the first term we observe that $L_1^h(x,t)$ is Lipschitz continuous in x. This is due to the fact that $G_0^h(x,x_k)$ is Lipschitz in x with a universal Lipschitz coefficient. Since the other terms in the expression of $L_1^h(x,t)$ in (3.20) are bounded by C, we have

$$|L_1^h(x,t)-L_2^h(y,t)|\le \frac{C}{h}|x-y|.$$

We use this fact and apply Lemma A.4 to deduce

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)(L_1^h(x,t)-L_1^h(y,t))\,dt\right)^4\le\frac{C}{h^4}|x-y|^4. \tag{3.31}$$

The constant C above depends on λ, Λ, and $\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}$.

To estimate the second term in (3.30), consider two distinct points $y<x$. Let j and k be the indices such that $x\in(x_{j-1},x_j]$ and $y\in(x_{k-1},x_k]$. Then one of the following holds: $j-k\ge 2$, $j-k=0$, or $j-k=1$. In the first case, since $|D^- U^0|\le Ch$ for some C depending on λ, Λ, and $\|f\|_2$, we have the following crude bound:

$$|L_2^h(x,t)-L_2^h(y,t)|\le C\le\frac{C}{h}|x-y|.$$

The same analysis leading to (3.31) applies, and the second term in (3.30) is bounded by $C|x-y|^4/h^4$ in this case.

When $|j-k|=0$, x and y are in the same interval (x_j, x_{j+1}). We can write

$$\int_0^1 q_\varepsilon(t)(L_2^h(x,t)-L_2^h(y,t))\,dt = \frac{a^* D^- U_{0j}^h}{h}\left(\int_y^x q_\varepsilon(t)\,dt - \frac{x-y}{h}\int_{I_j} q_\varepsilon(t)\,dt\right). \tag{3.32}$$

Since x and y are in the same interval, the function $(x-y)/h$ is bounded by one. Now Lemma A.4 applies, and we see that the fourth moments of the members in (3.32) are bounded by

$$C\left[\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_x^y q_\varepsilon(t)\,dt\right)^4 + \left(\frac{x-y}{h}\right)^4 \mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_{I_k} q_\varepsilon(t)\,dt\right)^4\right] \le C|x-y|^2.$$

When $j-k=1$, we have

$$\begin{aligned}\int_0^1 q_\varepsilon(t)L_2^h(y,t)\,dt &= \frac{a^* D^- U_{0j-1}^h}{h}\left(\int_{x_{j-2}}^y q_\varepsilon(t)\,dt - \frac{y-x_{j-2}}{h}\int_{x_{j-2}}^{x_{j-1}} q_\varepsilon(t)\,dt\right)\\ &= \frac{a^* D^- U_{0j-1}^h}{h}\left(-\int_y^{x_{j-1}} q_\varepsilon(t)\,dt - \frac{y-x_{j-1}}{h}\int_{x_{j-2}}^{x_{j-1}} q_\varepsilon(t)\,dt\right).\end{aligned}$$

Let x_{j-1} play the role of x in (3.32) and notice that $L_2^h(x_{j-1},t)=0$. We get

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(y,t)\,dt\right)^4 \le C\frac{|y-x_{j-1}|^2}{h^2}.$$

Similarly, in the interval where x lands, let x_{j-1} play the role of y in (3.32). We have

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(x,t)\,dt\right)^4 \le C\frac{|x-x_{j-1}|^2}{h^2}.$$

We combine these estimates and see that in this case, the second term in (3.30) is bounded by

$$\begin{aligned}&8\left(\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(y,t)\,dt\right)^4 + \mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)L_2^h(x,t)\,dt\right)^4\right)\\ &\quad \le C\frac{|x_{j-1}-y|^2+|x-x_{j-1}|^2}{h^2} \le C\frac{|x-y|^2}{h^2}.\end{aligned}$$

In the last inequality, we used the fact that $a^2+b^2 \le (a+b)^2$ for two nonnegative numbers a and b.

Combine these three cases to conclude that for any $x, y \in [0,1]$, the second term in (3.30) is bounded by $C|x-y|^2/h^2$. This, together with (3.31), shows

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_0^1 q_\varepsilon(t)(L^h(x,t)-L^h(y,t))\,dt\right)^4 \le C\frac{|x-y|^2}{h^4}. \tag{3.33}$$

In other words, $\{\varepsilon^{-\frac{1}{2}}I_\varepsilon(x)\}$ satisfies (3.28) with $\beta=4$ and $\delta=1$ and is therefore a tight sequence. Consequently, it converges to $\mathcal{I}^h$ in distribution in $\mathcal{C}$.

Convergence of $\{\varepsilon^{-\frac{1}{2}} r_\varepsilon^h\}$. For the convergence of finite-dimensional distributions, we need to show

$$\mathbb{E}\,\exp\left(i\cdot\frac{1}{\sqrt{\varepsilon}}\sum_{j=1}^{n}\xi^j r_\varepsilon^h(x^j)\right)\to 1$$

for any fixed n, $\{x^j\}_{j=1}^n$, and $\{\xi_j\}_{j=1}^n$. Since $|e^{i\theta}-1|\le|\theta|$ for any real number θ, the left-hand side of the equation above can be bounded by

$$\frac{1}{\sqrt{\varepsilon}}\mathbb{E}\left|\sum_j \xi_j r_\varepsilon^h(x_j)\right|\le\sum_j|\xi_j|\frac{1}{\sqrt{\varepsilon}}\sup_{1\le j\le n}\mathbb{E}|r_\varepsilon^h(x_j)|.$$

The last sum above converges to zero thanks to (3.19), completing the proof of convergence of finite-dimensional distributions.

For tightness, we recall that $r_\varepsilon^h(x)$ consists of r_{11}^h in (3.22), r_{12}^h in (3.23), $K_3(x)$ in (3.18), and $r_2^h(x)$ in (3.24). In the first three functions, x appears in Lipschitz continuous terms, e.g., in $D^-G_0^h(x;x_k)$ or $\phi_\varepsilon^j(x)-\phi_0^j(x)$. Meanwhile, the terms that are x-independent have mean square of order $O(\varepsilon)$ or less. Therefore, we can choose $\beta=2$ and $\delta=1$ in (3.28). For instance, we consider $r_{12}^h(x)$ in (3.23). Since q_ε is uniformly bounded, the integrals of q_ε on the interval I_k are bounded by Ch. Recall that $|D^-U_k^0|\le Ch$ also; we have

$$|r_{12}^h(x)-r_{12}^h(y)|\le\sum_{k=1}^{N}|D^-(G_0^h(x,x_k)-G_0^h(y,x_k))|\left(|\tilde r_{1k}|\int_{I_k}C|f|\,dt+C|\tilde r_{2k}|\right).$$

By the Lipschitz continuity of $D^-G_0^h$ and the estimate (3.17), we have

$$\mathbb{E}\left(\frac{r_{12}^h(x)-r_{12}^h(y)}{\sqrt{\varepsilon}}\right)^2\le C\frac{1}{\varepsilon}|x-y|^2\sup_k\mathbb{E}\{|\tilde r_{1k}|^2+|\tilde r_{2k}|^2\}\le C\frac{|x-y|^2}{h}.$$

Similarly, we can control r_{11}^h and K_3. For r_2^h in (3.24), we observe that it has the form of the main part of $K_2(x)$, which corresponds to $L_2^h(x,t)$ and the second term in (3.30), except the extra integral of q_ε. Therefore, the tightness argument for the second term in (3.30) can be repeated. The extra q_ε term is favorable: We can choose $\beta=2$ and $\delta=1$ in (3.28).

To summarize, $\{\varepsilon^{-\frac{1}{2}}r_\varepsilon^h/\sqrt{\varepsilon}\}$ can be shown to be tight by choosing $\beta=2$ and $\delta=1$ in (3.28). Therefore, it converges to the zero function in distribution in $\mathcal{C}$. We have thus established the convergence in (3.26).

The case of LRC. In this case, the scaling is $\varepsilon^{-\frac{\alpha}{2}}$. The proof is almost the same as above, and we only point out the key modifications.

Let us denote the right-hand side of (3.27) by $\mathcal{I}_H^h$. To show the convergence of the finite-dimensional distributions of $\{\varepsilon^{-\frac{\alpha}{2}}I_\varepsilon\}$, instead of using (3.29), we need the following analogue for random media with LRC:

$$\frac{1}{\varepsilon^{\frac{\alpha}{2}}}\int_0^1 q_\varepsilon(t)m(t)\,dt\xrightarrow[\varepsilon\to 0]{\text{distribution}}\sigma_H\int_0^1 m(t)\,dW_t^H, \tag{3.34}$$

where σ_H is defined below (2.19). The above holds only for $q(x,\omega)$ constructed as in (L1)–(L3), and $m\in L^1(\mathbb{R})\cap L^\infty(\mathbb{R})$. This convergence result was established in Theorem 3.1 of [3]. Assumptions on $q(x)$ in item (ii) allow us to use this result and conclude that the finite-dimensional distributions of $\{\varepsilon^{-\frac{\alpha}{2}} I_\varepsilon(x)\}$ converge to those of $\mathcal{I}_H^h$. For the tightness of $\{\varepsilon^{-\frac{\alpha}{2}} I_\varepsilon(x)\}$, we can follow the same procedures that lead to (3.31) and (3.32). We only need to consider second-order moments when applying the Kolmogorov criterion thanks to Lemma A.2, which says

$$\mathbb{E}\left(\frac{1}{\varepsilon^{\frac{\alpha}{2}}}\int_x^y q_\varepsilon(t)\,dt\right)^2 \leq C|x-y|^{2-\alpha}. \tag{3.35}$$

In the SRC case, since α equals one, we only have $|x-y|$ on the right. To get an extra exponent δ, we had to consider fourth moments. In the LRC case, α is less than one, so we gain a $\delta = 1-\alpha$ from the above estimate. With this in mind, we can simplify the proof we did for (3.26) to prove that $\{\varepsilon^{-\frac{\alpha}{2}} I_\varepsilon\}$ converges to $\mathcal{I}_H^h$. Similarly, $\{\varepsilon^{-\frac{\alpha}{2}} r_\varepsilon^h\}$ converges to the zero function in distribution, and hence in probability, in the space $\mathcal{C}$. The conclusion is that (3.27) holds. This completes the proof of Proposition 3.3. ∎

From the proofs of the propositions in this section, the results often hold if the conditions in item (i) or (ii) of Proposition 3.2 are violated in an ε-independent manner. For instance, if the second equation in (3.16) is modified to

$$b_\varepsilon^k - b_0^k = c(h)[1+\tilde{r}_{2k}]\left(-\frac{a^{*2}}{h^2}\int_{D_k} q_\varepsilon(t)\,dt\right) \tag{3.36}$$

for some function $c(h)$ and for region $D_k \subset I_k$, then this modification will be carried to $L^h(x,t)$ and following estimates, but the weak convergences in Proposition 3.3 still hold.

4. Weak convergence as *h* goes to 0. In the previous section, we established weak convergence of the corrector $u_\varepsilon^h - u_0^h$ of a general multiscale scheme when the correlation length ε of the random medium goes to zero while the discretization h is fixed. In this section, we send h to zero and characterize the limiting process. We aim to prove the following statement.

Proposition 4.1. *Let $L^h(x,t)$ be defined as in* (3.20). *As h goes to zero, the Gaussian processes on the right-hand sides of* (3.26) *and* (3.27) *have the following limits in distribution in $\mathcal{C}$:*

$$\sigma \int_0^1 L^h(x,t)\,dW_t \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}(x;W), \tag{4.1}$$

where $\mathcal{U}$ is the Gaussian process in (2.14). *Similarly,*

$$\sigma_H \int_0^1 L^h(x,t)\,dW_t^H \xrightarrow[h\to 0]{\text{distribution}} \mathcal{U}_H(x;W^H), \tag{4.2}$$

where $\mathcal{U}_H$ is the Gaussian process in (2.20).

We consider the case of SRC first. Recall that $\mathcal{I}^h(x)$ denotes the left-hand side of (4.1). It can be split further into three terms as follows. Let us first split $L_1^h(x,t)$ into two pieces:

$$L_{11}^h(x,t) = \sum_{k=1}^N \mathbf{1}_{I_k}(t)\frac{a^* D^- G_0^h(x,x_k)}{h}\cdot\frac{a^* D^- U_k^0}{h},$$

$$L_{12}^h(x,t) = \sum_{k=1}^N \mathbf{1}_{I_k}(t) a^* D^- G_0^h(x,x_k)\left(\frac{1}{h}\int_t^{x_k} f(s)\,ds - \frac{1}{h}\int_{x_{k-1}}^{x_k} f(s)\tilde{\phi}_0^k(s)\,ds\right). \tag{4.3}$$

Then define $\mathcal{I}_i^h(x)$ by

$$\mathcal{I}_i^h(x;W) = \sigma \int_0^1 L_{1i}^h(x,t)\,dW_t, \quad i=1,2; \quad \mathcal{I}_3^h(x;W) = \sigma \int_0^1 L_2^h(x,t)\,dW_t. \tag{4.4}$$

As it turns out, $\mathcal{I}_1^h(x;W)$ converges to the desired limit, while $\mathcal{I}_2^h(x;W)$ and $\mathcal{I}_3^h(x;W)$ converge to zero in probability.

Proof of (4.1) (convergence of $\{\mathcal{I}_1^h(x)\}$). By Proposition 3.4, we show the convergence of finite distributions of $\{\mathcal{I}_1^h(x)\}$ and tightness. Since all processes involved are Gaussian, for finite-dimensional distribution it suffices to consider the covariance function $R_1(x,y) \coloneqq \mathbb{E}\{\mathcal{I}_1^h(x)\mathcal{I}_1^h(y)\}$. By the Itô isometry of Wiener integrals, we have

$$R_1^h(x,y) = \sigma^2 \int_0^1 L_{11}^h(x,t)L_{11}^h(y,t)\,dt.$$

For any fixed x, $L_{11}^h(x,t)$, as a function of t, is a piecewise constant approximation of $L(x,t)$. This is obvious from the expression of $L(x,t)$ in (2.16). Therefore, $L_{11}^h(x,t)$ converges to $L(x,t)$ in (2.15) pointwise in t. Meanwhile, L_{11}^h is uniformly bounded as well. The dominant convergence theorem yields that for any x and y,

$$\lim_{h\to 0} R_1^h(x,y) = \sigma^2 \int_0^1 L(x,t)L(y,t)\,dt = \mathbb{E}(\mathcal{U}(x;W)\mathcal{U}(y;W)). \tag{4.5}$$

This proves convergence of finite-dimensional distributions.

The heart of the matter is to show that $\{\mathcal{U}_1^h(x;W)\}$ is a tight sequence. To this end, we consider its fourth moment

$$\mathbb{E}(\mathcal{I}_1^h(x) - \mathcal{I}_1^h(y))^4 = \int_{[0,1]^4} \prod_{i=1}^4 (L_{11}^h(x,t_i) - L_{11}^h(y,t_i))\mathbb{E}\prod_{i=1}^4 \mathrm{d}W_{t_i}. \tag{4.6}$$

Since increments in a Brownian motion are independent Gaussian random variables, we have

$$\mathbb{E}\prod_{i=1}^4 dW_{t_i} = [\delta(t_1-t_2)\delta(t_3-t_4) + \delta(t_1-t_3)\delta(t_2-t_4) + \delta(t_1-t_4)\delta(t_2-t_3)]\prod_{i=1}^4 dt_i. \tag{4.7}$$

Using this decomposition, and the fact that the L_{11}^h is piecewise constant, we rewrite the fourth moment above as three times of

$$\left(\int_0^1 (L_{11}^h(x,t) - L_{11}^h(y,t))^2\,dt\right)^2 = \left[\sum_{k=1}^N \left(\frac{a^* D^- (G_0^h(x,x_k) - G_0^h(y,x_k))}{h}\frac{a^* D^- U_k^0}{h}\right)^2 h\right]^2.$$

Hence, we need to control $\|L_{11}^h(x,\cdot) - L_{11}^h(y,\cdot)\|_2$. Since G_0^h is the Green's function associated with (1.2), it admits expression (2.17) as remarked below Proposition 3.2. Fix $y < x$, and let j_1 and j_2 be the indices so that $y \in (x_{j_1-1}, x_{j-1}]$ and $x \in (x_{j_2-1}, x_{j_2}]$. Then we can split the above sum into three parts. In the first part, k runs from one to $j_1 - 1$. In that case, both x_k and x_{k-1} are less than y. Formula (2.17) says $a^*(G_0^h(x,x_k) - G_0^h(y,x_k)) = x_k(y-x)$. Consequently,

$$\frac{a^* D^-(G_0^h(x,x_k) - G_0^h(y,x_k))}{h} = (y-x). \tag{4.8}$$

Since $|D^- U_k^0/h|$ is bounded, we have

$$\sum_{k=1}^{j_1-1} \left(\frac{a^* D^-(G_0^h(x,x_k) - G_0^h(y,x_k))}{h}\right)^2 \left(\frac{a^* D^- U_k^0}{h}\right)^2 h \le C|x-y|^2 \sum_{k=1}^{j_1-1} h \le C|x-y|^2. \tag{4.9}$$

Another part is k running from j_2+1 to N. In that case, both x_k and x_{k-1} are larger than x. The above analysis yields the same bound for this partial sum.

The remaining part is when k runs from j_1 to j_2. In this case, for some k, x_k may end up in (y,x), and we have to use different branches of (2.17) when evaluating $G_0^h(x,x_k)$ and $G_0^h(y,x_k)$. Consequently, the cancellation of h in (4.8) will not happen, and we need to modify our analysis. We observe that, due to the Lipschitz continuity of G_0^h and boundedness of $|D^- U^0/h|$, we always have

$$\sum_{k=j_1}^{j_2} \left(\frac{a^* D^-(G_0^h(x,x_k) - G_0^h(y,x_k))}{h}\right)^2 \left(\frac{a^* D^- U_k^0}{h}\right)^2 \cdot h \le C\frac{|x-y|^2}{h^2} \sum_{k=j_1}^{j_2} h. \tag{4.10}$$

If $j_2 - j_1 \le 1$, the last sum above is then bounded by $2C|x-y|^2/h$. In this case, it is clear that $|x-y| \le 2h$; as a result, the sum above is bounded by $C|x-y|$.

If $j_2 - j_1 \ge 2$, the above estimate will not help much if $j_2 - j_1$ is very large. Nevertheless, since $|D^- G_0^h/h|$ is bounded by some universal constant C, we have

$$\sum_{k=j_1}^{j_2} \left(\frac{a^* D^-(G_0^h(x,x_k) - G_0^h(y,x_k))}{h}\right)^2 \left(\frac{a^* D^- U_k^0}{h}\right)^2 \cdot h \le C \sum_{k=j_1}^{j_2} h. \tag{4.11}$$

Meanwhile, we observe that in this case

$$\begin{aligned} 3|x-y| &\ge 3(x_{j_2-1} - x_{j_1}) = 3(j_2 - j_1 - 1)h = (j_2 - j_1 + 1)h + 2(j_2 - j_1 - 2)h \\ &\ge (j_2 - j_1 + 1)h. \end{aligned}$$

Consequently, the sum in (4.11) is again bounded by $C|x-y|$. Combining these estimates, we have

$$\|L_{11}^h(x,\cdot) - L_{11}^h(y,\cdot)\|_2^2 \le C|x-y|. \tag{4.12}$$

It follows from the equation below (4.7) that $\{\mathcal{I}_1^h(x)\}$ is a tight sequence and hence converges to $\mathcal{U}(x,W)$.

Convergence of $\mathcal{I}_{12}^h$ to zero function. For the finite-dimensional distributions, we consider the covariance function $R_2^h(x,y) = \mathbb{E}\{\mathcal{I}_2^h(x)\mathcal{I}_2^h(y)\}$. By Itô isometry,

$$\sigma^2 \int_0^1 L_{12}^h(x,t) L_{12}^h(y,t)\, dt. \tag{4.13}$$

Now from the expression of $L_{12}^h(x,t)$, (4.3), we see that $L_{12}^h(x,t)$ converges to zero pointwise in t for any fixed x. Indeed, in the above expression, $|D^- G_0^h/h|$ is uniformly bounded, while the integrals of $f(s)$ and of $f(s)\tilde{\phi}_0^k(s)$ go to zero due to shrinking integration regions. Meanwhile, L_{12}^h is also uniformly bounded. The dominated convergence theorem shows $R_2(x,y) \to 0$ for any x and y, proving the convergence of finite-dimensional distributions. The tightness of $\{\mathcal{I}_2^h(x)\}$ is exactly the same as $\{\mathcal{I}_1^h(x)\}$; that is to say, the properties of $D^- G_0^h$ can still be applied. We conclude that $\{\mathcal{I}_2^h(x)\}$ converges to zero.

Convergence of $\mathcal{I}_3^h(x)$ *to zero.* For the finite-dimensional distributions, we observe that $L_2^h(x,t)$ is uniformly bounded, and for any fixed x, it converges to zero pointwise in t, due to shrinking of the nonzero interval $I_{j(x)}$. The covariance function of $\mathcal{I}_3^h(x)$, therefore, converges to zero, proving convergence of finite-dimensional distributions.

For tightness, we consider the fourth moment of $\mathcal{I}_3^h(x) - \mathcal{I}_3^h(y)$. By (4.7), it equals

$$\mathbb{E}(\mathcal{I}_3^h(x;W) - \mathcal{I}_3^h(y;W))^4 = 3\left(\int_0^1 (L_2^h(x,t) - L_2^h(y,t))^2 dt\right)^2. \tag{4.14}$$

Recalling the expression of $L_2^h(x,t)$ in (3.20), it is nonzero only on an interval of size h and is uniformly bounded. Let $j(x)$ be the interval where $L_2^h(x)$ is nonzero, and similarly define $j(y)$. Assume $y < x$ without loss of generality. Consider three cases: $j(x) = j(y)$, $j(y) = j(x) - 1$, and $j(x) - j(y) \geq 2$. In the first case, x and y fall in the same interval $[x_{j-1}, x_j]$ for some index j. Then we have

$$\int_0^1 (L_2^h(x,t) - L_2^h(y,t))^2 dt \leq C \int_0^1 \left(\mathbf{1}_{[x,y]}(t) - \frac{x-y}{h}\mathbf{1}_{I_j}(t)\right)^2 dt.$$

This integral can be calculated explicitly; it equals

$$\begin{aligned}&\int_0^1 \mathbf{1}_{[x,y]}(t) - 2\frac{x-y}{h}\mathbf{1}_{[x,y]} + \frac{(x-y)^2}{h^2}\mathbf{1}_{I_j}(t)dt\\ &\quad = (x-y) - 2\frac{x-y}{h}(x-y) + \frac{(x-y)^2}{h^2}h = (x-y)\left[1 - \frac{x-y}{h}\right].\end{aligned}$$

Since $|1-(x-y)/h| \leq 1$ and $|D^- U_k^0/h| \leq C$, the above quantity is bounded by $C|x-y|$.

In the second case, with j the unique index so that $y \leq x_j < x$ and using the triangle inequality, we have

$$\|L_2^h(x,t) - L_2^h(y,t)\|_2^2 \leq 2(\|L_2^h(x,t) - L_2^h(x_j,t)\|_2^2 + \|L_2^h(x_j,t) - L_2^h(y,t)\|_2^2).$$

For the first term of the right-hand side above, let x_j play the role of y in the previous calculation. This term is bounded by $C(x - x_j)$. Similarly, for the second term, let x_j play the role of x, and we bound this term by $C(x_j - y)$. Consequently, we can still bound $\|L_2^h(x,\cdot) - L_2^h(y,t)\|_2^2$ by $C|x-y|$.

In the third case, we have $h \leq |x-y|$. Meanwhile, since L_2^h is uniformly bounded and is nonzero only on intervals of size h, we have

$$\|L_2^h(x,t) - L_2^h(y,t)\|_2^2 \leq Ch \leq C|x-y|.$$

Combining these three cases, the conclusion is

$$\mathbb{E}(\mathcal{I}_3^h(x;W) - \mathcal{I}_3^h(y;W))^4 \leq C|x-y|^2. \tag{4.15}$$

This proves tightness and completes proof of the first item of Proposition 4.1. ∎

In the proof above, we used the fact that $G_0^h(x)$ defined in (3.21) is in fact the real Green's function defined in (2.17). However, the analysis follows as long as $|D_k^- G_0^h(x,x_k)/h|$ is piecewise Lipschitz in x with constant independent of h, and the total number of pieces does not depend on h.

The fact that $\mathcal{I}_2^h(x)$ and $\mathcal{I}_3^h(x)$ do not contribute to the limit is quite remarkable. It says the following. As long as the limiting distribution of the corrector $u_\varepsilon^h - u_0^h$ is considered, the role of the multiscale basis functions is mainly to construct the stiffness matrix, which is reflected by $\mathcal{I}_1^h(x)$; its roles in constructing the load vector F^ε and in assembling the global function, which are reflected in $\mathcal{I}_2^h(x)$ and $\mathcal{I}_3^h(x)$, respectively, are asymptotically not important.

Now, we prove the second part of Proposition 4.1.

Proof of (4.2). Recall that $\mathcal{I}_H^h(x)$ denotes the left-hand side of (4.2). Using the same splitting of L_1^h in (4.3), we can split $\mathcal{I}_H^h$ into three pieces $\mathcal{I}_{Hi}^h(x)$, $i = 1, 2, 3$, as in (4.4). The only necessary modification is to replace σ with σ_H and to replace the Brownian motion W_t with the fractional Brownian motion W_t^H. We show that $\mathcal{I}_{H1}^h(x)$ converges to $\mathcal{U}_H$ while $\mathcal{I}_{H2}^h$ and $\mathcal{I}_{H3}^h$ converge to the zero function.

Convergence of finite-dimensional distributions. For $\mathcal{I}_{H1}^h$, we consider the covariance matrix $R_{H1}^h(x,y)$ defined by $\mathbb{E}\{\mathcal{I}_{H1}^h(x)\mathcal{I}_{H1}^h(y)\}$. Using the isometry (A.19), we have

$$R_{H1}^h(x,y) = \kappa \int_0^1 \int_0^1 \frac{L_{11}^h(x,t) L_{11}^h(y,s)}{|t-s|^\alpha}\, dtds. \tag{4.16}$$

As before, the integrand in the above integral converges to $L(x,t)L(y,s)/|t-s|^\alpha$ for almost every (t,s). Meanwhile, since L_{11}^h is uniformly bounded, the integrand above is bounded by $C|t-s|^{-\alpha}$, which is integrable. The dominated convergence theorem then implies that R_{H1}^h converges to the covariance function of $\mathcal{U}_H(x; W^H)$. The convergence of finite distributions of $\mathcal{I}_{H2}^h$ and $\mathcal{I}_{H3}^h$ are similarly proved.

Tightness. Due to LRC, we only need to consider the second moments in (3.28). For $\{\mathcal{I}_{H1}^h\}$, we consider

$$\mathbb{E}(\mathcal{I}_{H1}^h(x) - \mathcal{I}_{H1}^h(y))^2 = \kappa \int_{\mathbb{R}^2} \frac{(L_{11}^h(x,t) - L_{11}^h(y,t))(L_{11}^h(x,s) - L_{11}^h(y,s))}{|t-s|^\alpha}\, dtds,$$

using again the isometry (A.19). Now we claim that

$$\|L_{11}^h(x,t) - L_{11}^h(y,t)\|_{L_t^p} \le C|x-y|^{\frac{1}{p}} \tag{4.17}$$

for any $p \ge 1$. Indeed, for $p = 2$, this is shown in (4.12); the analysis there actually shows also that the above holds for $p = 1$. For $p = \infty$, this follows from the uniform bound on L_{11}^h. For other p, this follows from interpolation; see [21, p. 75].

Now, we apply the Hardy–Littlewood–Sobolev lemma [21, section 4.3] to the expression of the second moment above. We obtain the bound

$$C(\alpha)\kappa \|L_{11}^h(x,\cdot) - L_{11}^h(y,\cdot)\|_{L^1} \|L_{11}^h(x,\cdot) - L_{11}^h(y,\cdot)\|_{L^{\frac{1}{1-\alpha}}} \le C|x-y|^{2-\alpha}.$$

Therefore, the Kolmogorov criterion (3.28) holds with $\beta = 2$ and $\delta = 1-\alpha$, proving tightness of $\{\mathcal{I}_{H1}^h\}$. Tightness of $\{\mathcal{I}_{H2}^h\}$ follows in the same way because L_{12}^h has the same structure as L_{11}^h as remarked before. Tightness of $\{\mathcal{I}_{H3}^h\}$ follows from the same argument above and the control on $\|L_2^h(x,\cdot) - L_2^h(y,\cdot)\|_2^2$ in the equation above (4.15). This completes the proof of (4.2). ∎

5. Applications to MsFEM in random media. In this section, we prove Theorems 2.1 and 2.2 as an application of the general results obtained in the preceding two sections by verifying that MsFEM satisfies the conditions of Proposition 3.3.

5.1. The multiscale basis functions. We describe MsFEM for (1.1) as a special case of the methods developed in [16], [17]. We verify that this scheme satisfies assumptions (N1)–(N3).

Recall that we have a uniform partition of the interval $[0,1]$ with nodal points $\{x_k\}_{k=1}^N$, where $x_0 = 0$ and $x_N = 1$, and I_k is the kth interval (x_{k-1}, x_k). The standard hat basis functions are denoted by $\{\phi_0^j(x)\}_{j=1}^{N-1}$. They span the space V_0^h. The idea of MsFEM is to replace the hat basis functions with $\{\phi_\varepsilon^j\}$, which are constructed as

$$\begin{cases} \mathcal{L}_\varepsilon \phi_\varepsilon^j(x) = 0, & x \in I_1 \cup I_2 \cup \cdots \cup I_{N-1}, \\ \phi_\varepsilon^j = \phi_0^j, & x \in \{x_k\}_{k=0}^N. \end{cases} \tag{5.1}$$

Here $\mathcal{L}_\varepsilon$ is the differential operator in (1.1). Clearly, ϕ_ε^j has the same support as ϕ_0^j and thus satisfies (N1). Note that the $\{\phi_\varepsilon^j\}$ are constructed locally on independent intervals and are suitable for parallel computing.

For any $k = 1, \ldots, N$, we observe that the only nonzero basis functions are ϕ_ε^k and ϕ_ε^{k-1}. Further, they sum up to one at the boundary points x_{k-1} and x_k. Since (5.1) is of linear divergence form, we conclude that $\phi_\varepsilon^k(x) + \phi_\varepsilon^{k-1}(x) \equiv 1$ on the interval. This shows that MsFEM satisfies (N3). In fact, the functions $\{\tilde{\phi}_\varepsilon^k\}_{k=1}^N$ for MsFEM are constructed by solving (5.1) on I_k with boundary values zero at x_{k-1} and one at x_k. Once they are constructed, $\{\phi_\varepsilon^j\}$ is given by (3.4). We can solve ϕ_ε^k analytically and obtain that

$$\tilde{\phi}_\varepsilon^j = b_\varepsilon^j \int_{x_{j-1}}^x a_\varepsilon^{-1}(t)\,dt, \qquad b_\varepsilon^j = \left(\int_{I_j} a_\varepsilon^{-1}(t)\,dt\right)^{-1}. \tag{5.2}$$

Consequently, (N1) and (N3) indicate that MsFEM also satisfies (N2). To calculate the entries of the stiffness matrix A_ε^h, we fix any $i = 2, \ldots, N-2$, and we compute

$$(A_\varepsilon^h)_{i-1i} = -\int_{I_i} a_\varepsilon \left(\frac{d\tilde{\phi}_\varepsilon^i}{dx}\right)^2 dx = -\left(a_\varepsilon \frac{d\tilde{\phi}_\varepsilon^i}{dx}\right)^2 \int_{I_i} a_\varepsilon^{-1}(s)\,ds = -b_\varepsilon^i.$$

The last equality can be verified from the fact that $\tilde{\phi}_\varepsilon^i$ solves (5.1) and integration by parts. For $i = 1$ and N, we verify that (3.3) holds for b_ε^0 and b_ε^N given by (5.2).

We record here a well-known superconvergence result: When dimension $d = 1$, the standard FEM is superconvergent in the sense that it yields exact values at nodal points. In our case, $u_0^h(x_k) = u_0(x_k)$, where u_0 solves (1.2) and u_0^h is the FEM approximation. We observe that this property is preserved by MsFEM. Indeed, let Pu_ε be the projection of u_ε in V_ε^h; i.e., $Pu_\varepsilon = u_\varepsilon(x_j)\phi_\varepsilon^j(x)$. Then, using integrations by parts, (5.1), and the fact that $Pu_\varepsilon - u_\varepsilon$ vanishes at nodal points, we have

$$A_\varepsilon(Pu_\varepsilon, v) = A_\varepsilon(u_\varepsilon, v) = F(v) \qquad \forall\ v \in V_\varepsilon^h.$$

Since the second equality is also satisfied by u_ε^h, it follows that $A_\varepsilon(Pu_\varepsilon - u_\varepsilon^h, v) = 0$ for any v in V_ε^h. In particular, by choosing $v = Pu_\varepsilon - u_\varepsilon^h$ and by coercivity of $A_\varepsilon(\cdot,\cdot)$, we conclude that $Pu_\varepsilon = u_\varepsilon^h$. The superconvergence result follows.

Several useful results follow from this superconvergent property. First, $u_\varepsilon^h(x)$ of MsFEM coincides with the true solution $u_\varepsilon(x)$ at nodal points. Note that u_ε can be explicitly solved analytically and that $|u_\varepsilon(x) - u_\varepsilon(y)| \le C|x-y|$ for some universal C. We then have

$$|D^- U_k^\varepsilon| = |u_\varepsilon^h(x_k) - u_\varepsilon^h(x_{k-1})| \le Ch. \tag{5.3}$$

This improves the condition (3.9) in Proposition 3.2 and hence improves several subsequent estimates. Second, a fact that we have used extensively before, we have $|D^- G_0^h| \le Ch$ and for any fixed x_k, $G_0^h(x;x_k)$ defined in (3.21) equals the continuous Green's function $G_0(x,x_k)$ for (1.2). This is because the functions agree at the nodal points due to superconvergence and they are both piecewise linear in x.

5.2. Proof of Theorems 2.1 and 2.2. Since MsFEM is a scheme that satisfies (N1)–(N3), in order to apply (3.12) and (3.20) in previous propositions, we only need to check that (3.16) and (3.17) hold.

Lemma 5.1. *Let $\tilde{\phi}_\varepsilon^k$ and b_ε^k be the functions in* (N1)–(N3) *for MsFEM defined in* (5.2). *Let $\tilde{\phi}_0^k$ and b_0^k be the corresponding functions for FEM.*

(i) *Suppose $a(x,\omega)$ and $q(x,\omega)$ satisfy* (S1)–(S3). *Then* (3.16) *and* (3.17) *hold and the conclusion of item* (i) *in Proposition* 3.2 *follows.*

(ii) *Suppose $a(x,\omega)$ and $q(x,\omega)$ satisfy* (L1)–(L3). *Then the conditions and hence the conclusions of item* (ii) *of Proposition* 3.2 *hold.*

Proof. From the explicit formulas (5.2), we have

$$b_\varepsilon^k - b_0^k = \left(\int_{I_k} \frac{1}{a_\varepsilon} dt\right)^{-1} - \left(\int_{I_k} \frac{1}{a^*} dt\right)^{-1} = -b_\varepsilon^k \frac{a^*}{h} \int_{I_k} q_\varepsilon(t) dt.$$

The above formula takes the form of the second line of (3.16) with

$$\tilde{r}_{2k} := -b_\varepsilon^k \int_{I_k} q_\varepsilon(t) dt. \tag{5.4}$$

Similarly, we have

$$\tilde{\phi}_\varepsilon^k(x) - \tilde{\phi}_0^k(x) = b_\varepsilon^k \left(\int_{x_{k-1}}^x q_\varepsilon(s) ds - \frac{x - x_{k-1}}{h} \int_{x_{k-1}}^{x_k} q_\varepsilon(s) ds \right). \tag{5.5}$$

This shows again that (3.16) holds with $\tilde{r}_{1k}$ having the same expression as $\tilde{r}_{2k}$ defined above. In (5.4), since $0 \le b_\varepsilon^k \le \Lambda h^{-1}$, we can apply Lemma A.1 in the case of SRC or apply Lemma A.2 in the case of LRC to conclude that $\mathbb{E}|\tilde{r}_{2k}|^2 \le Ch^{-1}\varepsilon$ in the first setting, while $\mathbb{E}|\tilde{r}_{2k}|^2 \le C(\varepsilon h^{-1})^\alpha$ in the second setting. This completes the proof. ∎

Note that the estimates (3.10) and (3.11) follow directly from this lemma. Therefore, we can apply Proposition 3.1 directly. Now we prove Theorem 2.1. Estimates (2.11) and (2.12) do not follow from Propositions 3.3 and 4.1 directly and need additional considerations.

Proof of Theorem 2.1 (finite element analysis). We have seen that u_ε^h superconverges to u_ε. From (1.1) and (5.1), we observe that the following equation holds on I_j for $j = 1, \ldots, N$:

$$\begin{cases} \mathcal{L}_\varepsilon(u_\varepsilon - u_\varepsilon^h) = f & \text{in } I_j, \\ u_\varepsilon^h - u_\varepsilon = 0 & \text{on } \partial I_j. \end{cases} \tag{5.6}$$

Using the ellipticity of the coefficient and integrations by parts, we obtain

$$\lambda|u_\varepsilon^h - u_\varepsilon|^2_{H^1,I_j} \le \int_{I_j} a_\varepsilon \frac{d}{dx}(u_\varepsilon^h - u_\varepsilon)\cdot\frac{d}{dx}(u_\varepsilon^h - u_\varepsilon)dx = \int_{I_j}(u_\varepsilon^h - u_\varepsilon)\mathcal{L}_\varepsilon(u_\varepsilon^h - u_\varepsilon)dx$$
$$= \int_{I_j} f(x)(u_\varepsilon^h - u_\varepsilon)(x)\,dx \le \|f\|_{2,I_j}\|u_\varepsilon^h - u_\varepsilon\|_{2,I_j}.$$

Now recall that the Poincaré–Friedrichs inequality says that

$$\|u_\varepsilon^h - u_\varepsilon\|_{2,I_j} \le \frac{h}{\pi}|u_\varepsilon^h - u_\varepsilon|_{H^1,I_j}. \tag{5.7}$$

Combining the inequalities above, we obtain

$$|u_\varepsilon^h - u_\varepsilon|_{H^1,I_j} \le \frac{h}{\lambda\pi}\|f\|_{2,I_j}.$$

Taking the sum over j, we obtain the first inequality in (2.11). To get the second inequality, we first apply the Poincaŕé–Friedrichs inequality to the equation above to get

$$\|u_\varepsilon^h - u_\varepsilon\|_{2,I_j} \le \frac{h^2}{\lambda\pi^2}\|f\|_{2,I_j} \tag{5.8}$$

and then sum over j. This completes the proof of (2.11) in item (i) of Theorem 2.1.

Energy norm of the corrector. By energy norm, we mean the $L^2(\Omega, L^2(D))$ norm. Recall the decomposition of the corrector into $K_i(x)$ in (3.15). For $K_1(x)$, we apply Cauchy–Schwarz to get the following bound for $|K_1|^2$:

$$\sum_i (U^\varepsilon - U^0)_i^2 \sum_j (\phi_0^j(x))^2 \le \sum_i (U^{\varepsilon_i} - U_i^0)^2\left(\sum_j \phi_0^j(x)\right)^2 = \|U^\varepsilon - U^0\|^2_{\ell^2}.$$

In the above derivation, we used the fact that $\phi_0^j(x)$ is nonnegative, and $\sum_j \phi_0^j(x) \equiv 1$. Now we apply (3.12) to control this term. The function $K_2(x)$, as in the proof of Proposition 3.2, can be written as $D^- U^0_{j(x)}(\tilde{\phi}_\varepsilon^{j(x)} - \tilde{\phi}_0^{j(x)})$. Then from (3.10), we have $\mathbb{E}|K_2(x)|^2 \le C\varepsilon\|R\|_{1,\mathbb{R}}$. For K_3, we have controlled $\mathbb{E}|K_3(x)|$ in (3.25). To control $\mathbb{E}|K_3(x)|^2$, we observe that $|K_3(x)| \le C\|f\|_2$. Note that all three estimates concluded in the three steps are uniform in x. Combining them, we complete the proof of (2.12).

Convergence in distribution as ε goes to zero. To prove item (ii) of Theorem 2.1, we apply (3.26) of Proposition 3.3. We need to verify (3.9) in addition to (3.16) and (3.17), which we already verified in the previous lemma. But this is implied by (5.3), and hence we obtain (2.13).

Convergence in distribution as h goes to zero. To prove (2.14), we apply the first result in Proposition 4.1. This completes the proof of the theorem. ∎

Proof of Theorem 2.2. In this case, the random processes $q(x)$ and $a(x)$ are constructed by (L1)–(L3). To prove the estimate in the energy norm, we follow the same steps as in the proof above but use item (ii) of Proposition 3.1 to control the term $\|U^\varepsilon - U^0\|^2_{\ell^2}$ in $K_1(x)$ and use Lemma A.2 to control the terms in $K_2(x)$ and $K_3(x)$.

To obtain the results in (2.19) and (2.20), we verify the conditions in item (ii) of Propositions 3.3 and 4.1, applying the second case in Lemma 5.1 and following the steps in the previous proof. This completes the proof of the theorem. ∎

6. Applications to HMM in random media. In this section, we adapt the general approach described in sections 3 and 4 to the case of HMM.

6.1. The heterogeneous multiscale method. The goal of HMM is to approximate the large-scale properties of the solution to (1.1) without computing the homogenized coefficient first. Suppose we already know this effective coefficient, i.e., a^* in our case. Then the large-scale solution u_0 can be solved by minimizing the functional

$$I[u] := \frac{1}{2}A_0(u,u) - F(u) = \frac{1}{2}\int_0^1 a^*\left(\frac{du}{dx}\right)^2 dx - \int_0^1 fu dx.$$

In numerical methods, the first integral above can be computed by the following midpoint quadrature rule:

$$A_0(u,u) \approx \sum_{j=1}^{N}\left(a^*(x^j)\frac{du}{dx}(x^j)\right)^2 h.$$

Here $x^j = (x_{j-1}+x_j)/2$ is the midpoint of I_j. In HMM, a^* is unknown, and the idea is to approximate $(u'a^*u')(x^j)$ by averaging in a local patch around the point x^j. For instance, we can take

$$\left(a^*(x^j)\frac{du}{dx}(x^j)\right)^2 \approx \frac{1}{\delta}\int_{I_j^\delta}\left(a_\varepsilon(s)\frac{d(\mathcal{L}u)}{dx}(s)\right)^2 ds.$$

Here, I_j^δ denotes the interval $x^j + \frac{\delta}{2}(-1,1)$, that is, the small interval centered in I_j with length δ. The operator $\mathcal{L}$ maps a function w in V_0^h, i.e., the space spanned by hat functions, to the solution of the following equation:

$$\begin{cases} \mathcal{L}_\varepsilon(\mathcal{L}w) = 0, & x \in I_1^\delta \cup \cdots \cup I_{N-1}^\delta, \\ \mathcal{L}w = w, & x \in \{\partial I_j^\delta\}_{j=1}^{N-1}. \end{cases} \tag{6.1}$$

The idea here is to encode small-scale structures of the random media into the construction of the bilinear form. The key difference that distinguishes HMM and MsFEM is that the above equations are solved for HMM on patches I_k^δ that are smaller than I_k. We check that $\mathcal{L}$ is a linear operator; therefore, the following approximation of $A_0(\cdot,\cdot)$ is indeed bilinear:

$$A_\varepsilon^\delta(w,v) := \sum_{j=1}^{N}\frac{h}{\delta}\int_{I_j^\delta} a_\varepsilon \frac{d(\mathcal{L}w)}{dx}\frac{d(\mathcal{L}v)}{dx}dx. \tag{6.2}$$

With this approximation of the bilinear form, HMM consists of finding

$$u_\varepsilon^{h,\delta} := \arg\min_{w\in V_0^h} \frac{1}{2}A_\varepsilon^\delta(w,w) - F(w).$$

This variational problem is equivalent to solving $u_\varepsilon^{h,\delta} = U_j^{\varepsilon,\delta}\phi_0^j(x)$, where $U^{\varepsilon,\delta}$ is determined by the linear system

$$A_\varepsilon^{h,\delta}U^{\varepsilon,\delta} = F^0. \tag{6.3}$$

Therefore, the above HMM can be viewed as an FEM. The finite-dimensional space here is V_0^h. Therefore HMM clearly satisfies (N1) and (N3). To check (N2), we calculate the associated stiffness matrix $A_\varepsilon^{h,\delta}$. It has entries $A_\varepsilon^\delta(\phi_0^i,\phi_0^j)$. From the defining equation (6.1), we see that $\mathcal{L}\phi_0^i$ is nonzero only on $I_i^\delta \cup I_{i+1}^\delta$, which implies that $A_\varepsilon^{h,\delta}$ is again tridiagonal. Further, we verify that $\mathcal{L}\phi_0^i + \mathcal{L}\phi_0^{i-1} = 1$ on the interval I_i^δ, which can be obtained from integrations by parts and which implies that $A_\varepsilon^{h,\delta}$ satisfies (3.3). Therefore, HMM satisfies (N2).

In fact, we can calculate the b_ε vectors. Let us consider the $(i-1,i)$th entry of $A_\varepsilon^{h,\delta}$, where i can be $2, \dots, N-1$. Since $(\mathcal{L}\phi_0^{i-1})' = -(\mathcal{L}\phi_0^i)'$ on I_i^δ, we have

$$(A^{h,\delta})_{\varepsilon i-1 i} = -\frac{h}{\delta}\int_{I_i^\delta} a_\varepsilon(s)\left(\frac{d(\mathcal{L}\phi^i)}{dx}\right)^2 ds.$$

Now from (6.1), we verify that $a_\varepsilon(\mathcal{L}\phi_0^i)'$ on I_i^δ is a constant given by

$$c_i^\delta = \left(\int_{I_i^\delta} a_\varepsilon^{-1}(s)\,ds\right)^{-1} \mathcal{L}\phi_0^i\Big|_{x^i-\frac{\delta}{2}}^{x^i+\frac{\delta}{2}} = \left(\int_{I_i^\delta} a_\varepsilon^{-1}(s)\,ds\right)^{-1}\frac{\delta}{h}.$$

Therefore, we have

$$(A^{h,\delta})_{\varepsilon i-1 i} = -(c_i^\delta)^2\frac{h}{\delta}\int_{I_i^\delta} a_\varepsilon^{-1}\,ds = -\frac{\delta}{h}\left(\int_{I_i^\delta} a_\varepsilon^{-1}(s)\,ds\right)^{-1} =: -b_{\varepsilon,\delta}^i. \tag{6.4}$$

We extend the definition of $b_{\varepsilon,\delta}^i$ to the cases of $i=1$ and $i=N$ and check that the $(1,1)$th and $(N-1,N-1)$th entries of the stiffness matrix also satisfy (3.3). In particular, the action of $A_\varepsilon^{h,\delta}$ on a vector satisfies the conservative form as in (3.5). In the following, to simplify notation, we drop the δ in the notation $A_\varepsilon^{h,\delta}$, $U^{\varepsilon,\delta}$, and $b_{\varepsilon,\delta}^i$.

The well-posedness of the optimization problem above, or equivalently of the linear system (6.3), is obtained by Lax–Milgram. We show that the bilinear form $A_\varepsilon^\delta(\cdot,\cdot)$ is continuous and coercive. Consider two arbitrary functions $w = W_i\phi_0^i$ and $v = V_j\phi_0^j$ in V_0^h. Then,

$$A_\varepsilon^h(w,v) = W_i A_{\varepsilon ij}^h V_j = -\sum_i W_i D^+(b_\varepsilon^i D^- V_i) = \sum_i D^- W_i b_\varepsilon^i D^- V_i.$$

Estimating the entries of vector b_ε by its infinity norm and using Cauchy–Schwarz, we obtain

$$|A_\varepsilon^h(w,v)| \le (\sup_{1\le i\le N} b_\varepsilon^i)\|D^- W\|_{\ell^2}\|D^- V\|_{\ell^2} \le \Lambda |w|_{H^1}|v|_{H^1}.$$

In the last inequality above, we used the fact that $\lambda h^{-1} \le b_\varepsilon^i \le \Lambda h^{-1}$, which can be seen from its definition in (6.4) and the uniform ellipticity of a_ε, and that $\|D^- W\|_{\ell^2} = |w|_{H^1}\sqrt{h}$ for $w \in V_0^h$. This proves continuity. Taking $w = v$, we have

$$A_\varepsilon^h(w,w) \ge (\inf_{1\le i\le N} b_\varepsilon^i)\|D^- W\|_{\ell^2}\|D^- W\|_{\ell^2} \ge \lambda |w|_{H^1}^2.$$

This proves coercivity. Therefore, by the Lax–Milgram theorem for the finite element space [24, p. 137], there exists a unique $u_\varepsilon^{h,\delta} \in V_0^h$ that solves the optimization problem. Further, we have

$$|u_\varepsilon^{h,\delta}|_{H^1} \le \frac{1}{\lambda}\sup_{w\in V_0^h}\frac{F(w)}{|w|_{H^1}} \le \frac{1}{\pi\lambda}\|f\|_2. \tag{6.5}$$

An immediate consequence is that $|D^- U^\varepsilon| \le C\sqrt{h}$ by the argument following Proposition 3.1.

6.2. Proof of Theorem 2.3. To prove Theorem 2.3, we apply Proposition 3.2 to write the corrector $u_\varepsilon^{h,\delta} - u_0^h$ as an oscillatory integral plus a lower order term. To apply Propositions 3.3 and 4.1 and obtain the weak convergences, we need to consider the difference $b_\varepsilon^k - b_0^k$ since $\tilde{\phi}_\varepsilon^j = \tilde{\phi}_0^j$ in HMM. From the expression of b_ε^k in (6.4), we have

$$b_\varepsilon^k - b_0^k = -b_\varepsilon^k \frac{a^*}{\delta}\int_{I_k^\delta} q_\varepsilon(t)\,dt = -(1+\tilde{r}_{2k})\frac{h}{\delta}\frac{a^{*2}}{h^2}\int_{I_k^\delta} q_\varepsilon(t)\,dt,$$

where $\tilde{r}_{2k}$ is a random variable defined by

$$\tilde{r}_{2k} = -\frac{h}{\delta}\,b_\varepsilon^k \int_{I_k^\delta} q_\varepsilon(t)\,dt.$$

We verify that in the case of SRC, i.e., when $q(x)$ satisfies (S1)–(S3), we have

$$\mathbb{E}|\tilde{r}_{2k}|^2 \le C\frac{\varepsilon}{\delta}\|R\|_{1,\mathbb{R}}, \qquad \mathbb{E}(b_\varepsilon^k - b_0^k)^2 \le C\frac{\varepsilon}{h^2\delta}\|R\|_{1,\mathbb{R}} \tag{6.6}$$

for some universal constant C. Comparing this with (3.11) and (3.17), we observe that the estimates have been multiplied by a factor $\frac{h}{\delta}$ in the HMM case. Similarly, it can be checked that in the case of LRC, i.e., when $q(x)$ satisfies (L1)–(L2), these estimates will be multiplied by a factor of $(\frac{\delta}{h})^\alpha$. With these formulas at hand, we prove the third main theorem of the paper.

Proof of Theorem 2.3 (short-range media and amplification effect). In this case, the difference of $b_\varepsilon^k - b_0^k$ and an estimate of it were captured in (6.6) and the equation above it. We cannot apply Propositions 3.2 and 3.3 directly. However, as remarked at the end of section 3.3, similar conclusions still hold. The same procedure as in the proof of Proposition 3.2 shows that the $L^h(x,t)$ function for HMM is

$$L^{h,\delta}(x,t) = \frac{h}{\delta}\sum_{k=1}^N \mathbf{1}_{I_k^\delta}(t)\,\frac{a^* D^- G_0^h(x,x_k)}{h}\,\frac{a^* D^- U_k^0}{h}. \tag{6.7}$$

The first weak convergence in (2.21) holds with this definition of $L^{h,\delta}$ as an application of a modified version of Proposition 3.3. Indeed, the proof there works with $L^{h,\delta}$ playing the role of L_{11}^h. The tightness is still obtained from the function $D^- G_0^h$, and the factor $\frac{h}{\delta}$ does not play any role at this stage.

When h goes to zero, we can follow the proof of Proposition 4.1 to verify the second convergence in (2.21). Indeed, tightness can be proved in exactly the same way. All that needs to be modified is the limit of the covariance function of $\mathcal{U}^{h,\delta}(x;W)$, which is defined to be $\sigma\int_0^1 L^{h,\delta}(x,t)\,dW_t$. This covariance function, by the Itô isometry, is as follows:

$$\begin{aligned} R^{h,\delta}(x,y) &:= \sigma^2\int_0^1 L^{h,\delta}(x,t)L^{h,\delta}(y,t)\,dt \\ &= \sigma^2\frac{h^2}{\delta^2}\sum_{k=1}^N \delta\,\frac{a^* D^- G_0^h(x;x_k)}{h}\,\frac{a^* D^- G_0^h(y;x_k)}{h}\left(\frac{a^* D^- U_k^0}{h}\right)^2. \end{aligned} \tag{6.8}$$

Recall the expression of L_{11}^h in (4.3). We verify that the above quantity can be written as

$$\sigma^2 \frac{h}{\delta} \int_0^1 L^h_{11}(x,t) L^h_{11}(y,t)\, dt.$$

Now the convergence in (4.5) implies that $R^{h,\delta}$ converges to the covariance function of $\sqrt{\frac{h}{\delta}}\mathcal{U}(x; W)$. This completes the proof of (2.21).

Long-range media. The expression for $b^k_\varepsilon - b^k_0$ is given above. Therefore, we can apply Propositions 3.2 and 3.3 (with modifications) to show that as ε goes to zero while h is fixed, the HMM corrector indeed converges to $\mathcal{U}^{h,\delta}_H(x; W^H)$ defined in (2.22). When h is sent to zero, we can follow the proof of Proposition 4.1 and show that $\mathcal{U}^{h,\delta}_H$ converges in distribution to some Gaussian process. To find its expression, we calculate the covariance function of $\mathcal{U}^{h,\delta}_H$. Thanks to the isometry (A.19), it is given by

$$R^{h,\delta}_H(x,y) := \kappa \int_0^1 \int_0^1 \frac{L^{h,\delta}(x,t) L^{h,\delta}(y,s)}{|t-s|^\alpha}\, dt ds. \tag{6.9}$$

Using the expression of $L^{h,\delta}$ and the short-hand notation

$$J_k(x) := \frac{a^* D^- G^h_0(x; x_k)}{h} \frac{a^* D^- U^0_k}{h},$$

the covariance function can be written as

$$\frac{\kappa h^2}{\delta^2}\left(\sum_{k=1}^N \sum_{m=1}^k [J_k(x)J_m(y) + J_m(x)J_k(y)] \int_{I^\delta_k}\int_{I^\delta_m} \frac{dtds}{|t-s|^\alpha} + \sum_{k=1}^N J_k(x)J_k(y) \int_{I^\delta_k}\int_{I^\delta_k} \frac{dtds}{|t-s|^\alpha}\right).$$

The integral of $|t-s|^{-\alpha}$ can be evaluated explicitly:

$$\frac{\kappa}{(1-\alpha)(2-\alpha)} \sum_{k=1}^N \sum_{m=1}^{k-1} [J_k(x)J_m(y) + J_m(x)J_k(y)] \frac{h^2}{\delta^2} ([(k-m)h + \delta]^{2-\alpha}$$
$$- 2[(k-m)h]^{2-\alpha} + [(k-m)h - \delta]^{2-\alpha}) + \frac{\kappa}{(1-\alpha)(2-\alpha)} \sum_{k=1}^N 2 J_k(x) J_k(y) \frac{h^2}{\delta^2} \delta^{2-\alpha}.$$

When $m < k$, the quantity between parentheses, together with the δ^2 on the denominator, forms a centered difference approximation of the second-order derivative of the function $r^{2-\alpha}$, evaluated at $(k-m)h$, i.e., at $t-s$. This derivative is precisely $(1-\alpha)(2-\alpha)|t-s|^{-\alpha}$. Meanwhile, the h^2 on the nominator can be viewed as the size of the measure $dtds$ on each block $I_k \times I_m$. Furthermore, $J_k(x)$ is precisely $L^h_{11}(x,t)$ evaluated on I_k. The conclusion is as follows: Those terms in the above equation with $m < k$ form an approximation of

$$\kappa \int_0^1 \int_0^1 \frac{L^h_{11}(x,t) L^h_{11}(y,s)}{|t-s|^\alpha}\, dtds.$$

The second sum corresponds to the diagonal terms $k = m$. Since $|J_k|$ is bounded, this sum is of order $O(h\delta^{-\alpha})$ and does not contribute in the limit as $h \to 0$, as long as $\delta \gg h^{\frac{1}{\alpha}}$. $R^{h,\delta}_H$ converges to the covariance function of $\mathcal{U}_H(x; W^H)$, finishing the proof of (2.22). ∎

7. A hybrid scheme that passes the corrector test. We now present a method that eliminates the amplification effect of HMM with $\delta < h$ exhibited in item (i) of Theorem 2.3 when the random media has SRC. Such an effect arises because the fluctuation in the short-range averaging effects occurring on the interval of size h is not properly captured by averaging occurring on an interval of size $\delta < h$.

The main idea is to subdivide the element I_k uniformly into M smaller patches and perform M independent calculations on each of these patches. This is a hybrid method that captures the idea of performing calculations on small intervals of size $\delta \ll h$ to reduce cost as in HMM, while preserving the averaging property of MsFEM by solving the elliptic equation on the whole domain.

Let $\delta = h/M$ be the size of the small patch I_k^ℓ for $1 \le \ell \le M$. Define $b_{\varepsilon\ell}^k$ by

$$b_{\varepsilon\ell}^k = \left(\frac{\delta}{h}\right)^2 \left(\int_{I_k^\ell} a_\varepsilon^{-1}(s)\,ds\right)^{-1}. \tag{7.1}$$

This definition is motivated by (6.4). Given a function w in the space V_0^h, we define its local projection into the space of oscillatory functions in the small patches I_k^ℓ by

$$\begin{cases} \mathcal{L}_\varepsilon(w_\ell^k) = 0, & x \in \cup_{k=1}^N \cup_{\ell=1}^M I_k^\ell, \\ w_\ell^k = w, & x \in \cup_{k=1}^N \cup_{\ell=1}^M \partial I_k^\ell, \end{cases} \tag{7.2}$$

where w_ℓ^k denotes this local projection. Recall that $\tilde{\phi}_0^k$ is the left piece of the hat basis function. Integrations by parts show that $b_{\varepsilon\ell}^k = A_\varepsilon((\tilde{\phi}_0^k)_\ell, (\tilde{\phi}_0^k)_\ell)$, where A_ε is the bilinear form defined in (3.2). HMM chooses one small patch $I_k^{\ell_*}$ and uses $A_\varepsilon((\tilde{\phi}_0^k)_{\ell_*}, (\tilde{\phi}_0^k)_{\ell_*}) = b_{\varepsilon\ell_*}^k$ to approximate the value $A_\varepsilon(\tilde{\phi}_0^k, \tilde{\phi}_0^k)$. Of course, the scaling h/δ is needed. This scaling factor turns out to amplify the variance as h goes to zero when the random medium has SRC.

We modify the method of HMM by constructing b_ε as $b_\varepsilon^k \coloneqq \sum_{\ell=1}^M b_{\varepsilon\ell}^k$. In other words, we define the stiffness matrix element $A_{\varepsilon i-1i}^h$ by $-\sum_{\ell=1}^M A_\varepsilon((\tilde{\phi}_0^i)_\ell, (\tilde{\phi}_0^i)_\ell)$. With this definition, we verify that

$$\begin{aligned} b_\varepsilon^k - b_0^k &= \sum_{\ell=1}^M \left(\frac{\delta}{h}\right)^2 \left[\left(\int_{I_k^\ell} a_\varepsilon^{-1}\,ds\right)^{-1} - \left(\int_{I_k^\ell} a^{*-1}\,ds\right)^{-1}\right] \\ &= \sum_{\ell=1}^M \left(\frac{a^*}{h}\right)^2 \left[-\int_{I_k^\ell} q_\varepsilon(s)\,ds + \left(\int_{I_k^\ell} q_\varepsilon(s)\,ds\right)^2 \left(\int_{I_k^\ell} a_\varepsilon^{-1}\,ds\right)^{-1}\right]. \end{aligned}$$

Rewriting the sum of the first terms in the parentheses, we obtain

$$b_\varepsilon^k - b_0^k = -\left(\frac{a^*}{h}\right)^2 \int_{I_k} q_\varepsilon(s)\,ds + r_\varepsilon^k,$$

where r_ε^k accounts for the sum over the second terms in the parentheses. Clearly, $\mathbb{E}|r_\varepsilon^k| \le C\varepsilon(h\delta)^{-1}$. This decomposition of $b_\varepsilon - b_0$ and the estimate of r_ε^k show that we can apply Proposition 3.2 to obtain the decomposition of the corrector. The $L^h(x,t)$ function in this case will be $L_{11}^h(x,t)$ in (4.3). Then it follows from Propositions 3.3 and 4.1 that the corrector in this method converges to the right limit.

In this modified method, all the local information on I_k is used to construct b_ε^k as in MsFEM. The main advantage is that the computation on $\{I_k^\ell\}_{\ell=1}^M$ can be done in a parallel manner. The calculation in MsFEM performed on a whole domain of size h

is replaced by h/δ independent calculations. Accounting for the coupling between the h/δ subdomains is necessary in MsFEM. It is no longer necessary in the modified method, which significantly reduces its complexity.

8. Conclusions and perspectives. This paper analyzes the theory of random fluctuations for several multiscale algorithms such as MsFEM and HMM in the simplified setting of a one-dimensional elliptic equation. One of our main results is that MsFEM and HMM with $\delta = h$ correctly capture the random fluctuations beyond the homogenization limit when the random media satisfy the SRC assumptions or the LRC assumptions considered in this paper. Such schemes pass the corrector test we have introduced and are therefore likely to be reliable in the analysis and the assessment of uncertainty quantifications even in settings where a well-understood theory of random fluctuations is not available.

We have also shown that less expensive schemes such as HMM with $\delta < h$ still capture the random fluctuations in media with LRC but may fail to capture them in media with SRC. The reason is that media with LRC display self-similarities across scales. As a consequence, the properties of the random fluctuations are correctly captured at the macroscopic scale h. On the other hand, media with SRC display averaging effects at the scale $\varepsilon \ll h$. Numerical schemes thus need to solve the equation at the microscopic scale on the whole domain in order to capture such effects. This forces the choice $\delta = h$ in the standard HMM scheme and leads us to the modification of the MsFEM and HMM schemes presented in section 7 to pass the corrector test for both LRC and SRC media.

The random fluctuations considered here all have a Gaussian structure and follow from an application of the central limit theorem for SRC media and a different but similar averaging procedure for LRC. Random fluctuations corresponding to rare events have to be analyzed with different methods. We refer the reader to [4] for the one-dimensional elliptic equation in the setting of large deviation theory.

The theory presented in this paper is restricted to the one-dimensional case. The main reason is that the theory of random fluctuations is well understood only in one dimension; see [3], [6]. In this setting, the random elliptic solution can be written as a stochastic integral whose analysis follows from fairly standard techniques. Note that the discretized solution solves a discrete system whose inversion is not explicit. This explains in large part why the results presented in this paper are rather long and technical.

In higher dimensions, only specific equations are amenable to fluctuation analyses; see [2], [14]. For such equations, the homogenization limit is trivial. Unlike the case considered in this paper, purely deterministic discretizations and basis functions based on the unperturbed elliptic equation provide accurate numerical solutions. Such models thus do not capture the difficulties inherent to the simulation of elliptic equations. They do not allow us to analyze, for instance, the resonance effects observed for MsFEM schemes and the oversampling techniques designed to eliminate such effects [12], [16], [17]. The theory of fluctuations and its effect on multiscale algorithms remain an open problem in such a setting.

Even though the one-dimensional equation is of purely academic interest, the analyses presented in this paper quantify the influence of using less expensive numerical schemes on the accuracy of the numerical solutions. In the case of SRC, random fluctuations are captured only by fairly expensive schemes that solve the elliptic equation on the whole domain of interest. In the case of LRC, self-similarity properties allow us to use much coarser schemes without sacrificing accuracy. These guiding rules should prevail for a large class of equations, elliptic and nonelliptic, linear and nonlinear, and hopefully in practical settings where a well-controlled theory of random fluctuations is not available.

Appendix A. Several useful lemmas.

A.1. Two moments. We first record two key estimates on oscillatory integrals, which we have used already. The first one accounts for random media with SRC.

LEMMA A.1. *Let $q(x,\omega)$ be a mean-zero stationary random process with integrable correlation function $R(x)$. Let $[a,b]$ and $[c,d]$ be two intervals on $\mathbb{R}$ and assume $b-a \le d-c$. Then*

$$\left| \mathbb{E} \int_a^b \int_c^d q_\varepsilon(t) q_\varepsilon(s)\, dt ds \right| \le \varepsilon (b-a) \|R\|_{1,\mathbb{R}}. \tag{A.1}$$

Proof. Let T denote the expectation of the double integral. It has the following expression:

$$T = \int_a^b \int_c^d R\left(\frac{t-s}{\varepsilon}\right) dt ds = \int_{\mathbb{R}} \int_{\mathbb{R}} R\left(\frac{t-s}{\varepsilon}\right) \mathbf{1}_{[a,b]}(t) \mathbf{1}_{[c,d]}(s)\, dt ds.$$

We change variables by setting $t \to t$ and $(t-s)/\varepsilon \to s$. The Jacobian of this change of variables is ε. Then we have

$$|T| \le \varepsilon \int_{\mathbb{R}} \int_{\mathbb{R}} |R(s)| \mathbf{1}_{[a,b]}(t)\, dt ds = \varepsilon (b-a) \|R\|_{1,\mathbb{R}}.$$

This completes the proof. ∎

The second one accounts for a special family of random media with LRC. The proof is adapted from [3].

LEMMA A.2. *Let $q(x,\omega)$ be defined as in* (L1)–(L3). *Let F be a function in the space $L^\infty(\mathbb{R})$. Let (a,b) and (c,d) be two open intervals and assume $b-a \le d-c$. Then we have*

$$\left| \mathbb{E} \int_a^b \int_c^d q\left(\frac{t}{\varepsilon}\right) q\left(\frac{s}{\varepsilon}\right) F(t) F(s)\, dt ds \right| \le C \varepsilon^\alpha (b-a)(d-c)^{1-\alpha}. \tag{A.2}$$

The constant C above depends only on κ, α, and $\|F\|_{\infty,\mathbb{R}}$.

Proof. By the definition of the correlation function R, we have

$$\mathbb{E}\left\{ \frac{1}{\varepsilon^\alpha} \int_a^b \int_c^d q\left(\frac{t}{\varepsilon}\right) q\left(\frac{s}{\varepsilon}\right) F(t) F(s)\, dt ds \right\} = \int_{\mathbb{R}^2} \varepsilon^{-\alpha} R\left(\frac{t-s}{\varepsilon}\right) F(t) \mathbf{1}_{[a,b]}(t) F(s) \mathbf{1}_{[c,d]}(s)\, dt ds.$$

As shown in [3], $R(\tau)$ is asymptotically $\kappa\tau^{-\alpha}$ with κ defined in (2.8). We expect to replace R with $\kappa\tau^{-\alpha}$ in the limit. Therefore, let us consider the difference

$$\int_{\mathbb{R}^2} \left| \varepsilon^{-\alpha} R\left(\frac{t-s}{\varepsilon}\right) - \frac{\kappa}{|t-s|^\alpha} \right| |F(t)| \mathbf{1}_{[a,b]}(t) |F(s)| \mathbf{1}_{[c,d]}(s)\, dt ds.$$

By the asymptotic relation $R \sim \kappa\tau^{-\alpha}$, we have for any $\delta > 0$ the existence of T_δ such that $|R(\tau) - \kappa\tau^{-\alpha}| \le \delta\tau^{-\alpha}$. Accordingly, we decompose the domain of integration into three subdomains:

$$D_1 = \{(t,s) \in \mathbb{R}^2, |t-s| \le T_\delta \varepsilon\},$$
$$D_2 = \{(t,s) \in \mathbb{R}^2, T_\delta \varepsilon < |t-s| \le 1\},$$
$$D_3 = \{(t,s) \in \mathbb{R}^2 \times, 1 < |t-s|\}.$$

On the first domain, we have

$$\int_{D_1} \left| \varepsilon^{-\alpha} R\left(\frac{t-s}{\varepsilon}\right) - \frac{\kappa}{|t-s|^\alpha} \right| |F(t)|\mathbf{1}_{[a,b]}(t)|F(s)|\mathbf{1}_{[c,d]}(s)\,dtds$$
$$\le \int_{D_1} \left| \varepsilon^{-\alpha} R\left(\frac{t-s}{\varepsilon}\right)\right| |F_1(t)||F_2(s)|dtds + \int_{D_1} \left| \frac{\kappa}{|t-s|^\alpha}\right| |F_1(t)||F_2(s)|dtds.$$

Here and below, we use the shorthand notation $F_1(t) = F(t)\mathbf{1}_{[a,b]}(t)$ and $F_2(s) = F(s)\mathbf{1}_{[c,d]}(s)$. The above integrals are then bounded by

$$\varepsilon^{-\alpha}\|R\|_{\infty,\mathbb{R}} \int_a^b |F(t)| \int_{t-T_\delta\varepsilon}^{t+T_\delta\varepsilon} |F_2(s)|dsdt + \int_a^b |F(t)| \int_{-T_\delta\varepsilon}^{T_\delta\varepsilon} \kappa|s|^{-\alpha}|F_2(t-s)|dsdt$$
$$\le \|F\|^2_{\mathbb{R},\infty}\left(2T_\delta\|R\|_{\mathbb{R},\infty} + \frac{2\kappa T_\delta^{1-\alpha}}{1-\alpha}\right)\varepsilon^{1-\alpha}.$$

On domain D_2, we have

$$\int_{D_2} \left| \varepsilon^{-\alpha} R\left(\frac{t-s}{\varepsilon}\right) - \frac{\kappa}{|t-s|^\alpha} \right| |F_1(t)||F_2(s)|dtds \le \delta \int_{D_2} |t-s|^{-\alpha}|F_1(t)||F_2(s)|dtds$$
$$\le 2\delta \int_a^b |F(t)| \int_{T_\delta\varepsilon}^1 |s|^{-\alpha}|F_2(t-s)|dsdt \le \frac{2\delta\|F\|^2_{\infty,\mathbb{R}}}{1-\alpha}(1 + T_\delta^{1-\alpha}\varepsilon^{1-\alpha}).$$

On domain D_3, we can bound $|t-s|^{-\alpha}$ by one, and we have

$$\int_{D_3} \left| \varepsilon^{-\alpha} R\left(\frac{t-s}{\varepsilon}\right) - \frac{\kappa}{|t-s|^\alpha} \right| |F_1(t)||F_2(s)|dtds \le \delta \int_{D_3} |t-s|^{-\alpha}|F_1(t)||F_2(s)|dtds$$
$$\le 2\delta \int_a^b |F(t)| \int_{T_\delta\varepsilon}^1 |F(t-s)|dsdt \le 2\delta\|F\|^2_{\infty,\mathbb{R}}(1 + T_\delta\varepsilon).$$

Therefore, for some constant C that does not depend on ε or δ, we have

$$\limsup_{\varepsilon\to0} \varepsilon^{-\alpha}\left|\mathbb{E}\int_a^b\int_c^d q\left(\frac{t}{\varepsilon}\right)q\left(\frac{s}{\varepsilon}\right)F_1(t)F_2(s) - \int_{\mathbb{R}^2}\frac{\kappa}{|t-s|^\alpha}F_1(t)F_2(s)\right| \le C\|F\|^2_{\infty,\mathbb{R}}\delta.$$

Sending δ to zero, we see that

$$\text{(A.3)} \qquad \lim_{\varepsilon\to0}\varepsilon^{-\alpha}\mathbb{E}\int_a^b\int_c^d q\left(\frac{t}{\varepsilon}\right)q\left(\frac{s}{\varepsilon}\right)F_1(t)F_2(s)dtds = \int_{\mathbb{R}^2}\frac{\kappa}{|t-s|^\alpha}F_1(t)F_2(s)dtds.$$

Finally, from the Hardy–Littlewood–Sobolev inequality [21, section 4.3], we have

$$\left| \int_{\mathbb{R}^2} \frac{F_1(t)F_2(s)}{|t-s|^\alpha}\, dtds \right| \le C\|F_1\|_{1,\mathbb{R}} \|F_2\|_{(1-\alpha)^{-1},\mathbb{R}} \le C\|F\|_\infty^2 (b-a)(d-c)^{1-\alpha}. \tag{A.4}$$

This completes the proof. ∎

A.2. Moment bound for stochastic process. In this section we provide a bound for the fourth-order moment of $q(x,\omega)$ in terms of the L^1 norm of the ρ-mixing coefficient.

Let $\mathcal{P}$ be the set of all ways of choosing pairs of points in $\{1,2,3,4\}$; i.e.,

$$\mathcal{P} := \{p = \{\{p(1),p(2)\},\{p(3),p(4)\}\} | p(i) \in \{1,2,3,4\}\}. \tag{A.5}$$

There are $C_6^2 = 15$ elements in $\mathcal{P}$.

LEMMA A.3. *Let $q(x,\omega)$ be a stationary mean-zero stochastic process. Assume $\mathbb{E}|q(0)|^4$ is finite and $q(x,\omega)$ is ρ-mixing with mixing coefficient $\rho(r)$ that is decreasing in r. Then we have*

$$\left|\mathbb{E}\left\{\prod_{i=1}^4 q(x_i)\right\}\right| \le \mathbb{E}|q(0)|^4 \sum_{p\in\mathcal{P}} \rho^{\frac12}(|x_{p(1)} - x_{p(2)}|)\rho^{\frac12}(|x_{p(3)} - x_{p(4)}|). \tag{A.6}$$

Proof. Given four points $\{q(x_i)\}$, $i = 1, \ldots, 4$, we can draw six line segments joining them. Among these line segments there is one that has the shortest length. Rearranging the indices if necessary, we assume it is the one joining x_1 and x_2. Then set $A = \{x_1, x_2\}$ and $B = \{x_3, x_4\}$. Rearranging the indices among each set if necessary, we assume also that $d(A,B)$ is obtained by $|x_1 - x_3|$. Then by the definition of ρ-mixing, we have

$$\begin{aligned} &\left|\mathbb{E}\left\{\prod_{i=1}^4 q(x_i)\right\} - R(x_1 - x_2)R(x_3 - x_4)\right| \\ &\quad \le \mathrm{Var}\{q(x_1)q(x_2)\}^{\frac12}\mathrm{Var}\{q(x_3)q(x_4)\}^{\frac12}\rho(|x_1 - x_3|). \end{aligned} \tag{A.7}$$

We can bound $\mathrm{Var}\{q(x_1)q(x_2)\}$, and similarly the variance of $\mathrm{Var}\{q(x_3)q(x_4)\}$, from above by $(\mathbb{E}|q(x_1)|^4\mathbb{E}|q(x_2)|^4)^{1/2}$. Therefore, the expression above can be bounded by $\mathbb{E}|q(0)|^4\rho(|x_1 - x_3|)$.

Since ρ is decreasing and $|x_1 - x_3| \ge |x_1 - x_2|$, we also have

$$\left|\mathbb{E}\left\{\prod_{i=1}^4 q(x_i)\right\} - R(x_1 - x_2)R(x_3 - x_4)\right| \le \mathbb{E}|q(0)|^4\rho(|x_1 - x_2|). \tag{A.8}$$

Now observe that $\min\{a,b\} \le (ab)^{\frac12}$ for any two nonnegative real numbers a and b. Applying this observation to the bounds of the two inequalities above, and using the triangle inequality, we obtain

$$\left|\mathbb{E}\left\{\prod_{i=1}^4 q(x_i)\right\}\right| \le \left|R(x_1 - x_2)\right| \cdot |R(x_3 - x_4)| + \mathbb{E}|q(0)|^4\rho^{\frac12}(|x_1 - x_2|)\rho^{\frac12}(|x_1 - x_3|). \tag{A.9}$$

Using the definition of mixing again, we obtain that $|R(x_1 - x_2)| = |\mathbb{E}q(x_1)q(x_2)|$ is bounded by

$$\mathrm{Var}^{\frac{1}{2}}(q(x_1))\mathrm{Var}^{\frac{1}{2}}(q(x_2))\rho(|x_1-x_2|) \le (\mathbb{E}|q(0)|^4)^{\frac{1}{2}}\rho^{\frac{1}{2}}(|x_1-x_2|).$$

In the last step, we used the fact that $\rho \le \rho^{\frac{1}{2}}$ since ρ can always be chosen no larger than 1. We can bound $R(x_1 - x_3)$ in the same way. Therefore, we obtain

$$\left|\mathbb{E}\left\{\prod_{i=1}^4 q(x_i)\right\}\right| \le \mathbb{E}|q(0)|^4[\rho^{\frac{1}{2}}(|x_1-x_2|)\rho^{\frac{1}{2}}(|x_3-x_4|) + \rho^{\frac{1}{2}}(|x_1-x_2|)\rho^{\frac{1}{2}}(|x_1-x_3|)]. \tag{A.10}$$

This completes the proof. □

We now derive a bound for the fourth-order moment of oscillatory integrals of q_ε.

LEMMA A.4. *Let $q(x,\omega)$ satisfy the conditions in the previous lemma. Assume in addition that the mixing coefficient satisfies that $\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}$ is finite. Let (x, y) be an interval in $\mathbb{R}$. Then for any bounded function $m(t)$, we have*

$$\mathbb{E}\left(\frac{1}{\sqrt{\varepsilon}}\int_x^y q\left(\frac{t}{\varepsilon}\right)m(t)\,dt\right)^4 \le 60\mathbb{E}|q(0)|^4 \cdot \|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}^2\|m\|_\infty^4 \cdot |x-y|^2. \tag{A.11}$$

Proof. The left-hand side of the desired inequality is

$$I = \frac{1}{\varepsilon^2}\int_x^y\int_x^y\int_x^y\int_x^y \mathbb{E}\prod_{i=1}^4 q\left(\frac{t_i}{\varepsilon}\right)\prod_{i=1}^4 m(t_i)\,d[t_1t_2t_3t_4]. \tag{A.12}$$

Here and below, $d[t_1 \cdot t_4]$ is a shorthand notation for $dt_1 \ \cdots \ d_4$. Applying the preceding lemma, we have

$$I \le \frac{\mathbb{E}|q(0)|^4\|m\|_\infty^4}{\varepsilon^2}\sum_{p\in\mathcal{P}}\int_x^y\int_x^y\int_x^y\int_x^y \rho^{\frac{1}{2}}\left(\frac{t_{p(1)}-t_{p(2)}}{\varepsilon}\right)\rho^{\frac{1}{2}}\left(\frac{t_{p(3)}-t_{p(4)}}{\varepsilon}\right)d[t_{p(1)} \ \cdots \ t_{p(4)}].$$

Note that we did not write an absolute sign for the argument in the ρ functions. We assume ρ is extended to be defined on the whole $\mathbb{R}$ by letting $\rho(x) = \rho(|x|)$. There are 15 terms in the sum above that are estimated in the same manner. Let us look at one of them, with $p(1) = p(3) = 1$, $p(2) = 2$, and $p(4) = 3$. We perform the following change of variables:

$$\frac{t_1-t_2}{\varepsilon} \to t_2, \quad \frac{t_1-t_3}{\varepsilon} \to t_3, \quad t_1 \to t_1, \quad t_4 \to t_4.$$

The Jacobian resulting from this change of variable cancels ε^2 on the denominator. The integral becomes

$$\int_x^y dt_1\int_x^y dt_4\int_{\frac{t_1-y}{\varepsilon}}^{\frac{t_1-x}{\varepsilon}}\rho^{\frac{1}{2}}(t_2)\,dt_2\int_{\frac{t_1-y}{\varepsilon}}^{\frac{t_1-x}{\varepsilon}}\rho^{\frac{1}{2}}(t_3)\,dt_3. \tag{A.13}$$

This integral is finite and is bounded from above by

$$|x-y|^2\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}}^2 = 4|x-y|^2\|\rho^{\frac{1}{2}}\|_{1,\mathbb{R}_+}^2. \tag{A.14}$$

The other terms in the sum have the same bound. Hence we have

$$I \le \mathbb{E}|q(0)|^4 \times 15 \times 4|x-y|^2 \|\rho^{\frac{1}{2}}\|^2_{1,\mathbb{R}_+} \|m\|^4_\infty. \tag{A.15}$$

This verifies (A.11) and completes the proof. □

A.3. Fractional Brownian motion. For the convenience of the reader, we briefly review some essential properties of fractional Brownian motion (fBm) and the stochastic integral with an fBm integrator.

An fBm $W^H(t)$ with Hurst index H is a mean-zero Gaussian process with $W^H(0) = 0$, stationary increments, and H-self-similarity; that is, for $a > 0$,

$$\{W^H(at)\}_{t\in\mathbb{R}} \overset{\mathcal{D}}{=} \{a^H W^H(t)\}_{t\in\mathbb{R}}, \tag{A.16}$$

where $\overset{\mathcal{D}}{=}$ means the equality in the sense of finite-dimensional distributions. From this similarity relation, we deduce $\mathbb{E}[(W^H(t))^2] = |t|^{2H}\mathbb{E}[(W^H(1))^2]$. In particular, if $\mathbb{E}[(W^H(1))^2] = 1$, we say the fBm is standard. It follows from the stationarity of increments that the covariance function of $W^H(t)$ is given by

$$R^H(t,s) = \mathbb{E}\{W^H(t)W^H(s)\} = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |s-t|^{2H}). \tag{A.17}$$

When $H = 1/2$, the increments of W^H are independent and the fBm reduces to the usual Brownian motion. For $H \ne 1/2$, the increments are stationary but not independent.

Stochastic integrals with respect to fBm can be defined on many functional spaces. Note that $H = 1 - \frac{\alpha}{2}$ is in the interval $(\frac{1}{2}, 1)$. In this case, a convenient functional space to define the stochastic integral is

$$|\Gamma|^H = \left\{ f\colon \int_\mathbb{R}\int_\mathbb{R} |f(x)||f(y)||x-y|^{2(H-1)}\,dxdy < \infty \right\}. \tag{A.18}$$

It is easy to check, for instance, from the Hardy–Littlewood-Sobolev lemma [21, section 4.3], that $L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \subset L^{1/H} \subset |\Gamma|^H$. Stochastic integrals against fBm do not satisfy Itô isometry; instead, we have

$$\mathbb{E}\left\{ \int_\mathbb{R} f(t)\,dW^H_t \int_\mathbb{R} h(s)\,dW^H_s \right\} = H(2H-1)\int_{\mathbb{R}^2} \frac{f(t)h(s)}{|t-s|^{2(1-H)}}\,dtds. \tag{A.19}$$

For a nice review on the stochastic integral with respect to fBm, we refer the reader to [23].

Acknowledgments. The authors acknowledge very fruitful discussions with Weinan E on multiscale algorithms and thank the reviewers for a thorough reading of the manuscript and remarks that helped with the presentation of the results.

REFERENCES

[1] T. Arbogast, *Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems*, SIAM J. Numer. Anal., 42 (2004), pp. 576–598.

[2] G. Bal, *Central limits and homogenization in random media*, Multiscale Model. Simul., 7 (2008), pp. 677–702.

[3] G. Bal, J. Garnier, S. Motsch, and V. Perrier, *Random integrals and correctors in homogenization*, Asymptot. Anal., 59 (2008), pp. 1–26.

[4] G. Bal, R. Ghanem, and I. Langmore, *Large deviation theory for a homogenized and corrected elliptic ODE*, J. Differential Equations, 251 (2011), pp. 1864–1902.

[5] P. Billingsley, *Convergence of Probability Measures*, Wiley Series in Probability and Statistics: Probability and Statistics, 2nd ed., John Wiley & Sons, New York, 1999.

[6] A. Bourgeat and A. Piatnitski, *Estimates in probability of the residual between the random and the homogenized solutions of one-dimensional second-order operator*, Asymptot. Anal., 21 (1999), pp. 303–315.

[7] Z. Chen and T. Y. Savchuk, *Analysis of the multiscale finite element method for nonlinear and random homogenization problems*, SIAM J. Numer. Anal., 46 (2008), pp. 260–279.

[8] P. Doukhan, *Mixing*, Lecture Notes in Statistics 85, Springer-Verlag, New York, 1994.

[9] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden, *Heterogeneous multiscale methods: A review*, Commun. Comput. Phys., 2 (2007), pp. 367–450.

[10] W. E, P. Ming, and P. Zhang, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18 (2005), pp. 121–156.

[11] Y. Efendiev, *The Multiscale Finite Element Method (MsFEM) and Its Applications*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1999.

[12] Y. R. Efendiev, T. Y. Hou, and X.-H. Wu, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 888–910.

[13] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics 19, AMS, Providence, RI, 1998.

[14] R. Figari, E. Orlandi, and G. Papanicolaou, *Mean field and Gaussian approximation for partial differential equations with random coefficients*, SIAM J. Appl. Math., 42 (1982), pp. 1069–1077.

[15] T. Y. Hou, *Numerical approximations to multiscale solutions in partial differential equations*, in Frontiers in Numerical Analysis (Durham 2002), Universitext, Springer-Verlag, Berlin, 2003, pp. 241–301.

[16] 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.

[17] 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.

[18] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, New York, 1994.

[19] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics 113, 2nd ed., Springer-Verlag, New York, 1991.

[20] S. M. Kozlov, *The averaging of random operators*, Mat. Sb., 109 (1979), pp. 188–202, 327.

[21] E. H. Lieb and M. Loss, *Analysis*, Graduate Studies in Mathematics 14, 2nd ed., AMS, Providence, RI, 2001.

[22] G. C. Papanicolaou and S. R. S. Varadhan, *Boundary value problems with rapidly oscillating random coefficients*, in Random Fields, Vol. I, II (Esztergom, 1979), Colloq. Math. Soc. János Bolyai 27, North–Holland, Amsterdam, 1981, pp. 835–873.

[23] V. Pipiras and M. S. Taqqu, *Integration questions related to fractional Brownian motion*, Probab. Theory Related Fields, 118 (2000), pp. 251–291.

[24] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics 23, Springer-Verlag, Berlin, 1994.

[25] M. Reed and B. Simon, *Methods of Modern Mathematical Physics.* II. *Fourier Analysis, Self-Adjointness*, Academic Press, New York, 1975.