Corrector Analysis of a Heterogeneous Multi-scale Scheme for Elliptic Equations with Random Potential

Guillaume Bal* Wenjia Jing†

September 23, 2012

Abstract

This paper analyzes the random fluctuations obtained by a heterogeneous multi-scale first-order finite element method applied to solve elliptic equations with a random potential. Several multi-scale numerical algorithms have been shown to correctly capture the homogenized limit of solutions of elliptic equations with coefficients modeled as stationary and ergodic random fields. Because theoretical results are available in the continuum setting for such equations, we consider here the case of a second-order elliptic equations with random potential in two dimensions of space.

We show that the random fluctuations of such solutions are correctly estimated by the heterogeneous multi-scale algorithm when appropriate fine-scale problems are solved on subsets that cover the whole computational domain. However, when the fine-scale problems are solved over patches that do not cover the entire domain, the random fluctuations may or may not be estimated accurately. In the case of random potentials with short-range interactions, the variance of the random fluctuations is amplified as the inverse of the fraction of the medium covered by the patches. In the case of random potentials with long-range interactions, however, such an amplification does not occur and random fluctuations are correctly captured independent of the (macroscopic) size of the patches.

These results are consistent with those obtained in [8] for more general equations in the one-dimensional setting and provide indications on the loss in accuracy that results from using coarser, and hence computationally less intensive, algorithms.

Keywords: Equations with random coefficients, multi-scale finite element method, heterogeneous multi-scale method, corrector test, long-range correlations.

AMS subject classification (2010): 35R60, 65N30, 65C99

1 Introduction

Differential equations with highly oscillatory coefficients arise naturally in many areas of applied sciences. The microscopic details of such equations are difficult to compute. Nev-

*Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027. Email: gb2030@columbia.edu

†Département de Mathématiques et Applications, Ecole Normale Supérieure, 45 Rue d'Ulm, 75230 Paris Cedex 05, France. Email: wjing@dma.ens.fr

ertheless, when the heterogeneous medium has certain properties involving separation of scales, periodicity, or stationary ergodicity, homogenization theories have been developed and they provide macroscopic models for the heterogeneous equations; see e.g. [19, 21, 25]. Many multi-scale algorithms have been devised to capture as much of the microscopic scale as possible without solving all the details of the micro-structure [2, 15, 14, 18]. Such a scheme is viewed as correct if it can well approximate the macroscopic solution when the heterogeneous medium satisfies conditions for homogenization to happen. Homogenization theory thus serves as a benchmark which ensures that the multi-scale scheme performs well in controlled environments, with the ! hope that it will still perform well in non-controlled environments, for instance when ergodicity and stationarity assumptions are not valid.

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 [9, 23]. When this is relevant, another benchmark for multi-scale numerical schemes that addresses the limiting stochasticity of the solutions is plausible: One computes the limiting (probability) distribution of the random fluctuation given by the multi-scale algorithm in the limit that the correlation length of the medium tends to 0 while the discretization size h of the scheme is fixed. If this h-dependent distribution converges, as $h \to 0$, to the limiting distribution of the corrector of the continuous equation (before discretization), we deduce that the multi-scale algorithm asymptotically correctly captures the randomness in the solution and passes the random corrector test.

Such proposal requires a controlled environment in which the theory of correctors is available. We introduced and analyzed such a benchmark in [8] using an ODE model whose corrector theory was studied in [11, 7]. The main purpose of this paper is to provide and analyze another benchmark using a PDE model whose corrector theory was studied in [5, 16, 6]. This is necessary because many multi-scale schemes that are different in higher dimensions turn out to be equivalent for the aforementioned ODE model, e.g. those in [1] and [18]. In the rest of this introduction, we first review some main results in [8]. Then we introduce the results of the current paper that address the corrector test using an elliptic PDE with random potential.

1.1 Corrector test using an ODE with random elliptic coefficient

The corrector test is based on the homogenization and corrector theory of the following equation:

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

Here, the diffusion coefficient $a(\frac{x}{\varepsilon},\omega)$ is obtained by rescaling $a(x,\omega)$ which is a random process on some probability space $(\Omega,\mathcal{F},\mathbb{P})$. It is well known [21, 25] that (and this generalizes to higher dimensions as well) when $a(x,\omega)$ is stationary, ergodic, and uniformly elliptic, then the solution u_ε converges to the following homogenized equation with deterministic and constant coefficient:

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

In the one-dimensional case, the coefficient a^* is the harmonic mean of $a(x)$, i.e., the inverse of the expectation of a^{-1}. We denote by $q(x)$ the deviation of $1/a(x)$ from its mean $1/a^*$. The corrector theories for the limiting distribution of $u_\varepsilon - u_0$ were studied by [7, 11]. The results in these papers are represented in path (iii) of the diagram in Fig. 1. The limiting distribution showing at the lower-right corner depends on the de-correlation rate of $q(x)$. When q is strongly mixing with integrable mixing coefficient (see (2.3) below), then $\beta = 1$ and W^β is a standard Brownian motion multiplied by σ, a factor determined by the correlation function of q as detailed in (2.2) below. When q has a heavy tail (is long-range) in the sense of (L1-L3) in section 2, we should take $\beta = \alpha$, $\alpha < 1$ being defined in (2.4), and W^α is the fractional Brownian motion with Hurst index $1 - \frac{\alpha}{2}$ multiplied by certain factor. These convergence results are understood as convergence in distribution in the space of continuous paths $\mathcal{C}([0,1])$.

The corrector test for multi-scale numerical schemes is therefore the following: Let h be the discretization size and $u_\varepsilon^h(x)$ the solution to (1.1) yielded by the scheme. Let $u_0^h(x)$ be the solution yielded by the same scheme applied to (1.2). The discrete corrector is $u_\varepsilon^h - u_0^h$. According to the de-correlation property of $q(x)$, we choose ε^β and interpret W^β as before. We say that a numerical procedure is consistent with the corrector theory and that it passes the corrector test when the diagram in Fig. 1 commutes:

Figure 1: A diagram describing the corrector test with a random ODE.

$$
\begin{tikzcd}
\dfrac{u_\varepsilon^h - u_0^h}{\sqrt{\varepsilon^\beta}}(x,\omega) \arrow[r, "h\to 0", "(i)"'] \arrow[d, "\varepsilon\to 0"', "(ii)"] & \dfrac{u_\varepsilon - u_0}{\sqrt{\varepsilon^\beta}}(x,\omega) \arrow[d, "(iii)"', "\varepsilon\to 0"] \\
\displaystyle\int L^h(x,y)\,dW^\beta(y) \arrow[r, "h\to 0", "(iv)"'] & \displaystyle\int (a^*)^2 \frac{\partial G}{\partial y}(x,y)\frac{\partial u}{\partial y}(y)\,dW^\beta(y).
\end{tikzcd}
$$

More precisely, we need to characterize the intermediate limit in path (ii) which appears on the left of the diagram. In this step, h is fixed while the correlation length ε is sent to zero. The intermediate limit distribution is h-dependent. Very often, it can be described as a stochastic integral as shown and we need to determine the kernel function $L^h(x,y)$. Next, we need to verify the converge path (iv) which is taken as $h \to 0$. The numerical scheme is said to pass (or fail) the corrector test if this limit holds (or does not).

In [8], we considered a *Finite Element Method* (FEM) based scheme in the framework of *Heterogeneous Multiscale Methods* (HMM), which is a general methodology for designing sublinear algorithms for multi-scale problems by exploiting special features of the problem, e.g. scale separation [15]. The macro-solver of this FEM-HMM scheme uses the standard P1 element on a uniform grid of size h. The corresponding discrete bilinear form which approximates the continuous bilinear form associated to (1.1) is

$$\mathcal{A}^h(u^h, v^h) = \sum_{j=1}^N \frac{du^h}{dx}(x_j) a^* \frac{dv^h}{dx}(x_j) h \approx \int_0^1 \frac{du^h}{dx}(x) a^* \frac{dv^h}{dx}(x) dx =: \mathcal{A}(u^h, v^h). \tag{1.3}$$

Here, a simple middle-point quadrature is used for the integral and x_j, $j = 1, \cdots, N = 1/h$

are the evaluation points. Since the effective coefficient a^* is unknown apriori, the FEM-HMM scheme approximates the discrete integrand by

$$\frac{du^h}{dx}(x_j)a^*\frac{dv^h}{dx}(x_j) \approx \frac{1}{\delta}\int_{I_{j\delta}} \frac{d\tilde{u}^h}{dx}(x)a_\varepsilon(x)\frac{d\tilde{v}^h}{dx}(x)dx,$$

where $I_{j\delta} = (x_j - \delta/2, x_j + \delta/2)$ is a patch inside the discretization interval $I_j = (x_j - h/2, x_j + h/2)$; the functions $\tilde{u}^h$ and $\tilde{v}^h$ are given in terms of $\{\tilde{\phi}^j\}$ where $\{\phi^j\}$ are the nodal bases and $\{\tilde{\phi}^j\}$ are given by the micro-solver

$$\begin{cases} -\dfrac{d}{dx}a_\varepsilon(x)\dfrac{d}{dx}\tilde{\phi}_j(x) = 0, & x \in I_{j\delta}, \\ \tilde{\phi}_j(x) = \phi_j(x), & x \in \partial I_{j\delta}. \end{cases} \tag{1.4}$$

When $\delta = h$, this scheme coincides with those in [18, 1]. It is known that one can choose $\delta < h$ to greatly reduce computational cost while still approximating the macroscopic solution quite well [15].

The main result of [8] shows that the corrector test for the above FEM-HMM scheme depends on the correlation structure of the random media. More precisely, for a long range correlated media (L1-L3 in section 2.1), the scheme is robust for the corrector test: the final limit in path (*iv*) of the diagram in Fig.1 agrees with the theoretical Gaussian limit for all $\delta \le h$. For a short range correlated media (S1-S3 in section 2.1), however, this holds true only for $\delta = h$. The final limit for $\delta < h$ is an amplified version of the theoretical Gaussian limit with an amplification factor $(h/\delta)^{1/2}$, which shows that reducing the computational cost results in an amplification of the variance of the numerical calculations.

1.2 Corrector test using elliptic PDE with random potential

The main objective of this paper is to provide a two dimensional corrector test. Such a strategy generalizes to arbitrary space dimensions, although for concreteness, we concentrate on the two-dimensional setting. A full theory of random fluctuations for second order elliptic PDE with highly oscillating random diffusion coefficients in dimension higher than one remains open and we can not use it for the corrector test. Instead, we base the test on the following elliptic equation with random potential:

$$\begin{cases} -\Delta u_\varepsilon + (q_0 + q_\varepsilon)u_\varepsilon(x,\omega) = f, & x \in Y, \\ u_\varepsilon(x,\omega) = 0, & x \in \partial Y. \end{cases} \tag{1.5}$$

The coefficient in the potential term consists of a smooth varying function q_0, and a highly oscillatory random function $q(\varepsilon^{-1}x,\omega)$ denoted by $q_\varepsilon(x)$ for simplicity. The random field $q(x,\omega)$ is assumed to be stationary ergodic and mean-zero. When ε goes to zero, the solution u_ε converges in $L^2(\Omega \times Y)$ to the homogenized solution u_0 that solves

$$\begin{cases} -\Delta u_0 + q_0 u_0(x) = f, & x \in Y, \\ u_0(x) = 0, & x \in \partial Y. \end{cases} \tag{1.6}$$

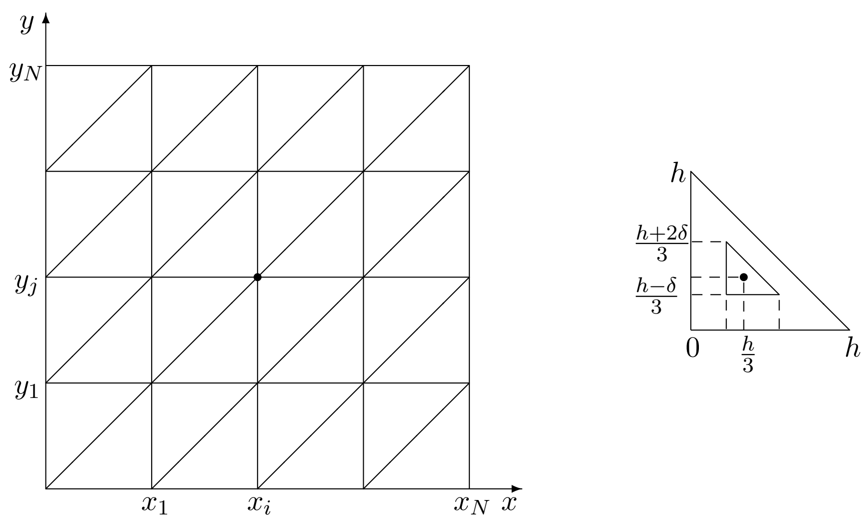
The corrector theory for the above homogenization is well understood; see [16, 5, 6]. When

Figure 2: A diagram describing the corrector test with a random PDE.

$$
\begin{tikzcd}
\left\langle \dfrac{u_\varepsilon^h - u_0^h}{\sqrt{\varepsilon^\beta}}(x,\omega), \varphi \right\rangle \arrow[r, "h\to 0", "(i)"'] \arrow[d, "\varepsilon\to 0"', "(ii)"] & \left\langle \dfrac{u_\varepsilon - u_0}{\sqrt{\varepsilon^\beta}}(x,\omega), \varphi \right\rangle \arrow[d, "(iii)"', "\varepsilon\to 0"] \\
\displaystyle\int_Y L^h[\phi](x,y)dW^\beta(y) \arrow[r, "h\to 0", "(iv)"'] & \displaystyle\int_Y \varphi(x)G(x,y)u_0(y)dW^\beta(y).
\end{tikzcd}
$$

the corrector $u_\varepsilon - u_0$ is properly scaled, it converges to a stochastic integral in a weak sense. This is described by the path (*iii*) of the diagram in Fig. 2. Both the scaling factor and the limit depend on the correlation structure of the random field. These results are reviewed in Section 2 below. As in the ODE (one-dimensional) setting, a corrector test can be sketched as in the diagram of Fig. 2. For a given multi-scale scheme, which yields u_ε^h and u_0^h when it is applied to (1.5) and (1.6), respectively, the main tasks are again to characterize the intermediate convergence in path (*ii*) while ε is sent to zero first and to check the validity of path (*iv*) when h is sent to zero afterwards.

Figure 3: Left: Triangulation of the unit square. Right: Shrinking from K to K_δ with respect to the barycenter.



Now we introduce a heterogeneous multi-scale scheme for (1.5). The weak formulation of the equation is to find u_ε in the Sobolev space $H_0^1(Y)$ so that $\mathcal{A}_\varepsilon(u_\varepsilon, v) = \langle f, v\rangle$ for all $v \in H_0^1(Y)$. Here and below, $\langle \cdot, \cdot \rangle$ denotes the usual pairing; $\mathcal{A}_\varepsilon$ is the bilinear form

$$\mathcal{A}_\varepsilon(u,v) = \int_Y \nabla u \cdot \nabla v + (q_0 + q_\varepsilon)uv \; dx, \quad \forall u, v \in H_0^1(Y). \tag{1.7}$$

Since we always assume that $q_0 + q_\varepsilon$ is positive, the weak formulation is well-posed thanks to the Lax-Milgram lemma. The scheme that will be considered is based on FEM. For simplicity, Y is taken as the two dimensional unit square $(0,1)^2$. Let $\mathcal{T}_h$ be the standard uniform triangulation as illustrated in Fig. 3. Here, the typical length of the triangles is $h = 1/N$ and N is the number of partitions on the axes. We consider first-order Lagrange elements. Associated to each (interior) nodal point (ih, jh), there is a continuous function ϕ^{ij} which is linear polynomial restricted to each triangle $K \in \mathcal{T}_h$ and which has value one at this nodal point and has value zero at all other nodal points. Note that the index i, j runs from 1 to $N-1$. The space V^h spanned by $\{\phi^{ij}\}$ is a finite dimensional subspace of $H^1_0(Y)$. The heterogeneous multi-scale scheme for (1.5) is to find $u^{h,\delta}_\varepsilon \in V!^h$ that satisfies

$$\mathcal{A}^{h,\delta}_\varepsilon(u^{h,\delta}_\varepsilon, v^h) = \langle f, v^h \rangle, \qquad \text{for all } v^h \in V^h, \tag{1.8}$$

where $\mathcal{A}^{h,\delta}_\varepsilon$ is a bilinear form on $V^h \times V^h$ which approximates $\mathcal{A}_\varepsilon$ as follows:

$$\mathcal{A}^{h,\delta}_\varepsilon(u^h, v^h) := \sum_{K \in \mathcal{T}_h} |K| \left(\frac{1}{|K_\delta|} \int_{K_\delta} \nabla u^h \cdot \nabla v^h + (q_0 + q_\varepsilon) u^h v^h \; dx \right). \tag{1.9}$$

Here, $K_\delta \subset K$ is a patch centered at the barycenter of K and has typical length δ; the symbol $|\cdot|$ means taking the area. $\mathcal{A}^{h,\delta}_\varepsilon$ hence is a numerical quadrature for the integral in (1.7) using averaged value around the barycenters of the elements. The scheme (1.8) is analyzed in Section 3 and it is well-posed.

When the above scheme is applied to the homogenized equation (1.6), it yields a solution $u^{h,\delta}_0$ in V^h so that

$$\mathcal{A}^{h,\delta}_0(u^{h,\delta}_0, v^h) = \langle u^{h,\delta}_0, v^h \rangle, \qquad \text{for all } v^h \in V^h, \tag{1.10}$$

and $\mathcal{A}^{h,\delta}_0$ is given by

$$\mathcal{A}^{h,\delta}_0(u^h, v^h) := \sum_{K \in \mathcal{T}_h} |K| \left(\frac{1}{|K_\delta|} \int_{K_\delta} \nabla u^h \cdot \nabla v^h + q_0 u^h v^h \; dx \right).$$

The discrete corrector function is defined to be the difference between $u^{h,\delta}_\varepsilon$ and $u^{h,\delta}_0$.

Remark 1.1. The ratio $|K_\delta|/|K|$ is the two dimensional analog of δ/h in the aforementioned FEM-HMM scheme for the ODE setting. It measures savings in the computational cost. As in the ODE setting, we expect the corrector test to depend on the ratios $\{|K_\delta|/|K| \,:\, K \in \mathcal{T}_h\}$. To simplify notations, we assume $|K_\delta|/|K|$ is a constant for all K. For instance, consider a typical triangle K with vertices $(0,0), (h,0)$ and $(0,h)$. K_δ can be obtained by shrinking with respect to the barycenter $(h/3, h/3)$ so that it has vertices $((h-\delta)/3, (h-\delta)/3), ((h+2\delta)/3, (h-\delta)/3)$ and $((h-\delta)/3, (h+2\delta)/3)$; see Fig. 3. If all K_δ are obtained in this manner, we have $\sqrt{|K_\delta|/|K|} = (\delta/h)^{d/2}$ with $d = 2$. More general patches than those of the paper could also be considered without changing our main conclusions; we concentrate on a specific construction of the sma! ll patches K_δ to simplify notation and denote the ratio $\sqrt{|K_\delta|/|K|}$ by δ/h.

1.3 Main Results

The main results of this paper concern the limiting distribution of the discrete corrector $u^{h,\delta}_\varepsilon - u^{h,\delta}_0$ with proper scaling. They depend on the correlation structure of the random field q_ε. We refer to section 2.1 below for notation. In particular, SRC (respectively LRC) stands for short (respectively long) range correlation.

Theorem 1.2. *Let $u_\varepsilon^{h,\delta}$ and $u_0^{h,\delta}$ be the solutions obtained from the heterogeneous multi-scale schemes* (1.8) *and* (1.10)*, respectively. Assume that q_0 is a nonnegative constant and f is in $\mathcal{C}^2(\overline{Y})$, i.e., twice continuously differentiable over $\overline{Y}$. For an arbitrary test function $\varphi \in \mathcal{C}^2(\overline{Y})$, the following holds.*

(1) *In the SRC setting, as ε goes to zero while h and δ satisfying $h > \delta \gg \varepsilon$ are fixed, we have*

$$\frac{1}{\sqrt{\varepsilon^d}} \int_Y \varphi(x)[u_\varepsilon^{h,\delta} - u_0^{h,\delta}]dx \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sigma \int_Y L^h[\varphi](x)dW(x). \tag{1.11}$$

Here $L^h[\varphi]$ is a bounded function defined in (4.19) *below.*

(2) *As h and δ go to zero with the ratio $\delta/h < 1$ being fixed, we have*

$$\sigma \int_Y L^h[\varphi](x)dW(x) \xrightarrow[h \to 0]{\text{distribution}} \frac{h}{\delta}\sigma \int_Y \mathcal{G}\varphi(x)u_0(x)dW(x). \tag{1.12}$$

(3) *In the LRC setting, the above convergences are modified by*

$$\frac{1}{\sqrt{\varepsilon^\alpha}} \int_Y \varphi(x)[u_\varepsilon^{h,\delta} - u_0^{h,\delta}]dx \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sqrt{\kappa} \int_Y L^h[\varphi](x)W^\alpha(dx), \tag{1.13}$$

and

$$\sqrt{\kappa} \int_Y L^h[\varphi](x)W^\alpha(dx) \xrightarrow[h \to 0]{\text{distribution}} \sqrt{\kappa} \int_Y \mathcal{G}\varphi(x)u_0(x)W^\alpha(dx). \tag{1.14}$$

Above, $\mathcal{G}$ is the solution operator of (1.6)*; κ is defined after* (2.5) *below. W is the standard multi-parameter Wiener process; $W^\alpha(dy)$ is formally defined to be $\widetilde{W}^\alpha(y)dy$ and $\widetilde{W}^\alpha(y)$ is a Gaussian random field with covariance function given by $\mathbb{E}\{\widetilde{W}^\alpha(x)\widetilde{W}^\alpha(y)\} = |x-y|^{-\alpha}$.*

Remark 1.3. We refer the reader to [20] for theories of stochastic integrals with respect to multi-parameter random processes. In fact, the limits above can be written as the following Gaussian distributions:

$$\sigma \int_Y \mathcal{G}\varphi(x)u_0(x)dW(x) \quad \stackrel{\text{distribution}}{=} \mathcal{N}(0, \sigma^2\|u_0\mathcal{G}\varphi\|_{L^2}^2), \tag{1.15}$$

$$\sqrt{\kappa} \int_Y \mathcal{G}\varphi(x)u_0(x)W^\alpha(dx) \quad \stackrel{\text{distribution}}{=} \mathcal{N}(0, \int_{Y^2} \frac{\kappa(u_0\mathcal{G}\varphi) \otimes (u_0\mathcal{G}\varphi)}{|x-y|^\alpha}dxdy). \tag{1.16}$$

Comparing these results with Theorem 2.1 below recalling the theory of random fluctuations in the continuous setting, and with the paths in Fig. 2, we find in the LRC setting that the multi-scale scheme (1.8) captures the theoretical Gaussian limit fluctuations after ε and h are successively sent to zero. Furthermore, the scheme is robust in the sense that it provides the correct fluctuations for arbitrary small patches with $0 < \delta < h$ (both being independent of ε). For SRC medium, however, the correct limit for the random fluctuations is captured only when $\delta = h$, that is $K_\delta = K$ for all $K \in \mathcal{T}_h$. The amplification effect in the case of $\delta < h$ is again characterized by $(h/\delta)^{\frac{d}{2}}$. The main results hence generalize the findings of [8] to a higher dimensional setting.

Remark 1.4. The main results are stated under the assumptions in Remark 1.1. When the ratios $\{|K|/|K_\delta|\}$ are not uniform, the limit in (1.12) does not have a simple form and must account for the non-uniform amplification factors over different triangulation elements. Nevertheless, the main conclusions in the above result are not modified. This remark applies to the ODE setting in [8] also.

The rest of this paper is devoted to a proof of the main theorem. Preliminary material on random fields and the corrector theory in the continuous scale are provided in Section 2. Then main ingredient of the proof is a conservative structure of the stiffness matrix associated to the multi-scale scheme; this is considered in section 3. Similar structures have been observed and explored in other settings [18, 8]. It allows us to write the discrete corrector in the form of oscillatory random integrals. Their limiting distributions are then characterized using well established techniques in [16, 5, 6]. This is done in Section 4. These sections also include some useful results on the scheme, such as the H^1 estimate of the solution to (1.8), which are interesting in their own right. We conclude this introduction by several comments.

1.4 Further Discussions

This paper studies the specific multi-scale scheme (1.8) for the elliptic equation (1.5) with a random potential. The analysis takes advantage of the conservative structure of the stiffness matrix. We refer to Proposition 3.4 below for a detailed statement. Other schemes possessing this property can be analyzed similarly. To simplify the presentation, we considered first-order nodal basis on a uniform triangulation. For higher order schemes in which basis functions occupy larger sub-domain of Y, and for general regular triangulation where different nodal basis may occupy different number of triangles, the structure in the stiffness matrix is more complicated. Nevertheless, we believe that the analysis should extend without major differences to this more general setting.

The scheme (1.8) fits within the framework of HMM. We refer to [15] for references on this method applied to operators of the form $\mathcal{L}_\varepsilon = \sum_{\alpha,\beta=1}^d \partial_\alpha(a_{\alpha\beta}(x/\varepsilon)\partial_\beta)$, which are more complex to analyze than (1.5). In our scheme, the macro-solver is essentially a standard finite element method for the homogenized equation (1.6), and the micro-solver essentially estimates the effect of the random potential on small patches K_δ.

Other multiscale schemes and methodologies have been developed for $\mathcal{L}_\varepsilon$ using properties of the medium such as separation of scales, periodicity, or ergodicity, e.g. [3, 4, 18]. For instance, the *Multiscale Finite Element Method* (MsFEM) in [18] constructs oscillatory bases by solving $\mathcal{L}_\varepsilon$-problems on the supports of the nodal bases $\{\phi^{ij}\}$ and uses the so-called over-sampling strategy to diminish the resonance errors introduced by the artificial boundary conditions of the local $\mathcal{L}_\varepsilon$-problems. It would be interesting to investigate how random fluctuation are captured by this scheme and in particular what is the effect of the over-sampling strategy. The differential operator in (1.6) does not exhibit such resonances, and hence this paper does not address such issues.

Other multi-scale schemes approach differential operators with rough coefficients like $\mathcal{L}_\varepsilon$ without assuming any separation of scales or special properties. For instance, [24] constructs oscillatory bases by solving $\mathcal{L}_\varepsilon$-problems on sub-domains that are larger than the supports of $\{\phi^{ij}\}$ but still small compared to the whole domain Y. It was proved there, using the so-called transfer property of the divergence operator [10], that the resulting finite dimensional space can be used to solve the whole $\mathcal{L}_\varepsilon$-problem with errors that are independent of the regularity of $\{a_{\alpha\beta}\}$. Analyzing the fluctuations in such schemes is beyond the scope of this paper.

2 Review of Corrector Theory in the Continuous Scale

In this section, we review the corrector theories for (1.5) developed in [16, 5]. They are formulated for the following random fields.

2.1 Random field settings

In the elliptic equation (1.5), the heterogeneous potential, denoted by $\tilde{q}_\varepsilon(x)$ henceforth, consists of a slowly varying part $q_0(x)$ and a highly oscillating part $q_\varepsilon(x)$. The latter is modeled as $q(\frac{x}{\varepsilon}, \omega)$, that is, spatially rescaled from some random field $q(x, \omega)$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In the sequel, $\mathbb{E}$ denotes the mathematical expectation with respect to the probability measure $\mathbb{P}$.

We assume that $q(x, \omega)$ is *stationary.* That is to say, for any positive integer k and k-tuple $(x_1, \cdots, x_k)$, for any point z and any Borel measurable set $\mathcal{A} \subset \mathbb{R}^k$, one has

$$\mathbb{P}\{(q(x_1), \cdots, q(x_k)) \in \mathcal{A}\} = \mathbb{P}\{(q(x_1 + z), \cdots, q(x_k + z)) \in \mathcal{A}\}.$$

With this assumption, q admits an (auto-)correlation function $R(x)$ defined by

$$R(x) := \mathbb{E}q(y)q(y + x) = \mathbb{E}q(0)q(x). \tag{2.1}$$

It is easy to check that R is symmetric and positive definite. Due to Bochner's theorem [27], the Fourier transform of R is a positive Radon measure. In particular, when R is integrable, one can define

$$\sigma^2 := \int_{\mathbb{R}^d} R(x)dx, \tag{2.2}$$

and it is a finite non-negative number. Without loss of generality, we also assume that q is mean-zero.

A key parameter of the random field that will determine different limiting correctors is the de-correlation rate. It is an indicator of how fast (with respect to distance) the random field becomes independent.

Recall that a random field $q(x, \omega)$ is said to be ρ-mixing with mixing coefficient ρ if there exists some function $\rho(r)$, which maps $\mathbb{R}_+$ to $\mathbb{R}_+$ and vanishes as r tends to infinity, so that for any Borel sets $A, B \subset \mathbb{R}^d$, the sub-$\sigma$-algebras $\mathcal{F}_A$ and $\mathcal{F}_B$ generated by the process restricted on A and B respectively de-correlate rapidly as follows:

$$\sup_{\xi \in L^2(\mathcal{F}_A), \eta \in L^2(\mathcal{F}_B)} \left| \frac{\mathbb{E}\ \xi\eta - \mathbb{E}\xi\ \mathbb{E}\eta}{(\mathrm{Var}\ \xi\ \mathrm{Var}\ \eta)^{1/2}} \right| \le \rho(d(A, B)). \tag{2.3}$$

Here $d(A, B)$ is the distance between the sets A and B. The function ρ characterizes the decay of the dependence of the random field at different places. We refer the reader to [13] for more information on mixing properties of random fields.

We consider two settings of random fields. In the first case, we say that $q(x,\omega)$ is *short range correlated* (SRC). This means

(S1) q is ρ-mixing with mixing coefficient $\rho(r)$ such that $\rho(|x|) \in L^1(\mathbb{R}^d)$.

(S2) $|q(x)| \le C$ so that $\tilde{q}_\varepsilon(x)$ is positive for a.e. $\omega \in \Omega$.

(S3) In this case, the correlation function $R(x)$ is integrable over $\mathbb{R}^d$ and we assume that σ defined in (2.2) is positive.

In the second case, we say that $q(x,\omega)$ is *long range correlated* (LRC). In fact, we consider the very specific setting as follows.

(L1) $q(x)$ has the form $\Phi \circ g(x)$, where $g(x,\omega)$ is a centered stationary Gaussian random field with unit variance and heavy tail, i.e.

$$R_g(x) := \mathbb{E}\{g(y)g(y+x)\} \sim \kappa_g |x|^{-\alpha} \quad \text{as } |x| \to \infty, \tag{2.4}$$

for some positive constant κ_g and some real number $\alpha < d$.

(L2) The function Φ is uniformly bounded so that $\tilde{q}_\varepsilon(x)$ is positive for a.e. ω. Further, we assume the Fourier transform $\hat{\Phi}$ satisfies that $\int |\hat{\Phi}|(1+|\xi|^3)$ is finite.

(L3) We assume also that Φ has Hermite rank one, that is

$$\int_{\mathbb{R}} \Phi(s) e^{-\frac{s^2}{2}} ds = 0, \quad V_1 := \int_{\mathbb{R}} s\Phi(s) e^{-\frac{s^2}{2}} ds \neq 0. \tag{2.5}$$

As a consequence $\kappa := V_1^2 \kappa_g$ defines a positive number. For more information on the Hermite rank, we refer the reader to [29].

2.2 Corrector theory in the continuous scale

The corrector theory for the elliptic equation with random potential, that is the limiting distribution of the difference between u_ε and u_0 which solve (1.5) and (1.6) respectively, has been investigated in [16, 5] in the SRC setting, and in [6] in the LRC setting. Using the notations and random field settings introduced above, the results in dimension two of these references can be summarized as follows.

Theorem 2.1 ([16, 5, 6])**.** *Let u_ε and u_0 be as above and let the dimension $d = 2$. Denote by $G(x,y)$ be the fundamental solution to the Dirichlet problem* (1.6). *When the random potential $q(x)$ satisfies the SRC setting, we have*

$$\frac{u_\varepsilon(x) - u_0(x)}{\sqrt{\varepsilon^d}} \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sigma \int_Y G(x,y) u_0(y) dW(y) \tag{2.6}$$

weakly in the spatial variable. When the random potential satisfies the LRC setting, we have

$$\frac{u_\varepsilon(x) - u_0(x)}{\sqrt{\varepsilon^\alpha}} \xrightarrow[\varepsilon \to 0]{\text{distribution}} \sqrt{\kappa} \int_Y G(x,y) u_0(y) W^\alpha(dy) \tag{2.7}$$

weakly in the spatial variable.

Here, W and W^α are the same as in Theorem 1.2. The convergences above are weakly in the spatial variable in the sense of (1.15) and (1.16).

3 Analysis of the Discrete Equation

In this section, we analyze the heterogeneous multi-scale scheme (1.8) in detail. In particular, we prove that the scheme with $\varepsilon \ll \delta \le h$ admits a unique solution in the space V^h that approximates u_0 in H^1. With the standard uniform triangulation, we show that the stiffness matrix associated to the scheme has some conservative form, which allows us to write the discrete corrector conveniently in terms of their coordinates. In the next section, we use this discrete representation to prove the main theorem.

3.1 Well-posedness of the scheme

The multi-scale scheme (1.8) with $\delta = h$ coincides with the standard FEM and is well-posed. For the sake of completeness, we show that this holds also for $\delta < h$.

Recall that V^h is the finite dimensional subspace of $H_0^1(Y)$ with nodal basis $\{\phi^{ij}\}$ defined in section 1.2. We have defined three quadratic forms: $\mathcal{A}_\varepsilon$ for the heterogeneous equation (1.5), $\mathcal{A}_\varepsilon^{h,\delta}$ for the heterogeneous multi-scale scheme which is an approximation of $\mathcal{A}_\varepsilon$ by local integration, and $\mathcal{A}_0^{h,\delta}$ which is like $\mathcal{A}_\varepsilon^{h,\delta}$ but uses the mean coefficient q_0 only. In fact, $\mathcal{A}_0^{h,\delta}$ may also be seen as a heterogeneous scheme for the homogenized equation (1.6), which in turn is associated with the quadratic form

$$\mathcal{A}_0(u,v) = \int_Y \nabla u \cdot \nabla v + q_0 uv \; dx, \quad u, v \in H_0^1(Y). \tag{3.1}$$

Let x_K denote the barycenter of the element $K \in \mathcal{T}_h$. Exploring the expression of $\mathcal{A}_\varepsilon^{h,\delta}$ and $\mathcal{A}_0^{h,\delta}$, we further define formally

$$\hat{\mathcal{A}}_\varepsilon^{h,\delta}(u^h, v^h)[x_K] = \fint_{K_\delta} \nabla u^h \cdot \nabla v^h + (q_0 + q_\varepsilon) u^h v^h \; dx,$$

and similarly define $\hat{\mathcal{A}}_0^{h,\delta}(u^h, v^h)[x_K]$. Hereafter, the integral symbol with a dash in the middle denotes the averaged integral.

We define the error due to the heterogeneous choice of integrating element (K_δ instead of K) by

$$e(\text{HMS}) := \max_{K \in \mathcal{T}_h} \sup_{V^h \ni u^h, v^h \neq 0} \frac{|K| |\hat{\mathcal{A}}_\varepsilon^{h,\delta}(u^h, v^h)[x_K] - \hat{\mathcal{A}}_0^{h,\delta}(u^h, v^h)[x_K]|}{\|u^h\|_{H^1(K)} \|v^h\|_{H^1(K)}}. \tag{3.2}$$

With this notation we have the following theorem.

Theorem 3.1. *Assume that q_0 is a nonnegative constant, and that $q_\varepsilon(x) + q_0$ is uniformly bounded and nonnegative. There exist unique solutions in V^h for the schemes* (1.8) *and* (1.10). *Let $u_\varepsilon^{h,\delta}$, $u_0^{h,\delta}$ be the solutions. Let u_0 solves* (1.6). *For $\varepsilon \ll \delta \le h$, we have*

$$\|u_\varepsilon^{h,\delta} - u_0\|_{H^1} \le C(h + e(\text{HMS})), \tag{3.3}$$

The above estimates hold also if we replace $u_\varepsilon^{h,\delta}$ by $u_0^{h,\delta}$ and delete the term $e(\text{HMS})$.

Proof. Let p be either ε or 0. The existence and uniqueness follow from

$$\mathcal{A}_p^{h,\delta}(v^h, v^h) \geq C\|v^h\|_{H^1}^2, \qquad \text{for any } v^h \in V^h.$$

Indeed, because ∇v^h is constant on $K \in \mathcal{T}_h$ and $q_0 + q_\varepsilon$ is non-negative, we have

$$\mathcal{A}_p^{h,\delta}(v^h, v^h) \geq \sum_{K\in\mathcal{T}_h} |K| \fint_{K_\delta} |\nabla v^h|^2 dx = \sum_{K\in\mathcal{T}_h} \int_K |\nabla v^h|^2 dx = |v^h|_{H^1}^2 \geq C\|v^h\|_{H^1}^2.$$

Here and in the sequel, $|\cdot|_{H^1}$ and $|\cdot|_{W^{k,p}}$ are the standard semi-norms of the corresponding Sobolev spaces.

We apply the first Strang lemma (Theorem 4.1.1 of [12]), and obtain

$$\|u_0 - u_\varepsilon^{h,\delta}\|_{H^1} \leq C \inf_{v^h\in V^h} \Big(\|u_0 - v^h\|_{H^1} + \sup_{w^h\in V^h} \frac{|\mathcal{A}_\varepsilon^{h,\delta}(v^h, w^h) - \mathcal{A}_0(v^h, w^h)|}{\|w^h\|_{H^1}} \Big).$$

Set $v^h = \Pi u_0$, the projection of u_0 to the space V^h. From classical interpolation result, e.g. Theorem 3.1.6 of [12], we have

$$\|\Pi u_0 - u_0\|_{H^1} \leq Ch\|u_0\|_{H^2}.$$

For any $w^h \in V^h$, we have

$$|\mathcal{A}_\varepsilon^{h,\delta}(v^h, w^h) - \mathcal{A}_0(v^h, w^h)| \leq |\mathcal{A}_\varepsilon^{h,\delta}(v^h, w^h) - \mathcal{A}_0^{h,\delta}(v^h, w^h)| + |\mathcal{A}_0^{h,\delta}(v^h, w^h) - \mathcal{A}_0(v^h, w^h)|.$$

For the first term, we have

$$\begin{aligned}
|\mathcal{A}_\varepsilon^{h,\delta}(v^h, w^h) - \mathcal{A}_0^{h,\delta}(v^h, w^h)| &\leq \sum_{K\in\mathcal{T}_h} |K| |\hat{\mathcal{A}}_\varepsilon^{h,\delta}(v^h, w^h)[x_K] - \hat{\mathcal{A}}_0^{h,\delta}(v^h, w^h)[x_K]| \\
&\leq e(\text{HMS}) \sum_{K\in\mathcal{T}_h} \|v^h\|_{H^1(K)} \|w^h\|_{H^1(K)} \\
&\leq e(\text{HMS}) \|v^h\|_{H^1} \|w^h\|_{H^1}.
\end{aligned}$$

In the equalities above, we used the definition of $e(\text{HMS})$ and Cauchy-Schwarz respectively. For the second term, we first observe that

$$\begin{aligned}
\mathcal{A}_0^{h,\delta}(v^h, w^h) - \mathcal{A}_0(v^h, w^h) = \sum_{K\in\mathcal{T}_h} \Bigg\{ \frac{|K|}{|K_\delta|} \left(\int_{K_\delta} q_0 v^h w^h \, dx - |K_\delta| (q_0 v^h w^h)(x_K) \right) \\
- \left(\int_K q_0 v^h w^h \, dx - |K| (q_0 v^h w^h)(x_K) \right) \Bigg\}.
\end{aligned}$$

The items in the sum can be recognized as errors of barycenter numerical approximation of integrals. Error estimate for such numerical quadrature is discussed in the next lemma and by (3.4) we have that $|\mathcal{A}_0^{h,\delta}(v^h, w^h) - \mathcal{A}_0(v^h, w^h)|$ is bounded by

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h} C\|q_0\|_{C^1} \left\{ \frac{h^2}{\delta^2} \delta \|v^h\|_{H^1(K_\delta)} \|w^h\|_{L^2(K_\delta)} + h\|v^h\|_{H^1(K)} \|w^h\|_{L^2(K)} \right\} \\
\leq& Ch\|q_0\|_{C^1} \sum_{K\in\mathcal{T}_h} \|v^h\|_{H^1(K)} \|w^h\|_{L^2(K)} \leq Ch\|q_0\|_{C^1} \|v^h\|_{H^1} \|w^h\|_{H^1}.
\end{aligned}$$

Combining the above estimates, we find that

$$\|u_0 - u_\varepsilon^{h,\delta}\|_{H^1} \le \|\Pi u_0 - u_0\|_{H^1} + (e(\mathrm{HMS}) + Ch)\|\Pi u_0\|_{H^1} \le C(h + e(\mathrm{HMS})).$$

The constant depends on $\|q_0\|_{C^1}$, $\|u_0\|_{H^1}$ and some uniform bound of h/δ and hence is independent of ε, h or δ. □

The following lemma concerns error estimate for barycenter numerical quadrature of product of two functions in $P_1(K)$, the space of linear polynomials on a triangular element K. It is stated in the simplest setting thought it can be generalized to regular element easily. This lemma is used in the proof of the previous theorem.

Lemma 3.2. *Let $\hat{K}$ be an isosceles right triangle with unit side length. Let K be the image of $\hat{K}$ under the linear transform $F(\hat{x}) = B\hat{x} + b \in \mathbb{R}^2$. For any function ψ on K, let $\hat{\psi}$ be the function $\psi \circ F$. Assume $q_0 \in W^{1,\infty}(K)$. Then for any $v, w \in P_1(K)$, we have*

$$\left| \int_K q_0(x)v(x)w(x)\, dx - |K|(q_0 v w)(x_K) \right| \le C\|B\|\|q_0\|_{W^{1,\infty}(K)}\|v\|_{H^1(K)}\|w\|_{L^2(K)}. \quad (3.4)$$

Here, x_K is the barycenter of K; $\|B\|$ is the matrix norm of B.

Proof. We follow the steps in the proof of [12, Theorem 4.1.4]. Consider any $\psi \in W^{1,\infty}(K)$ so that $\hat{\psi} = \psi \circ F$ is in $W^{1,\infty}(\hat{K})$. Let $|E(\psi w)|$ denote the error of the barycenter quadrature for the integral $\int_K \psi w dx$. After change of variables,

$$E(\psi w) = |\det(B)| \left(\int_{\hat{K}} \hat{\psi}(\hat{x})\hat{w}(\hat{x})\, d\hat{x} - |\hat{K}|(\hat{\psi}\hat{w})(\hat{x}_{\hat{K}}) \right) = |\det(B)|\hat{E}(\hat{\psi}\hat{w}).$$

On the reference element $\hat{K}$, since all norms on $P_1(\hat{K})$ are equivalent, we have

$$|\hat{E}(\hat{\psi}\hat{w})| \le \hat{C}\|\hat{\psi}\|_{L^\infty(\hat{K})}\|\hat{w}\|_{L^\infty(\hat{K})} \le \hat{C}\|\hat{\psi}\|_{W^{1,\infty}(\hat{K})}\|\hat{w}\|_{L^2(\hat{K})}.$$

We view $\hat{E}(\cdot\ \hat{w}) : \hat{\psi} \mapsto \hat{E}(\hat{\psi}\hat{w})$ as a linear functional on $W^{1,\infty}(\hat{K})$. The above estimate shows that $\hat{E}(\cdot\ \hat{w})$ is continuous with norm less than $\hat{C}\|\hat{w}\|_{L^2(\hat{K})}$. We check also that $\hat{E}(\cdot\ \hat{w})$ vanishes on $P_0(\hat{K})$, the space of constant functions on $\hat{K}$. Therefore, due to Bramble-Hilbert lemma [12, Theorem 4.1.3], there exists some $\hat{C}$ such that for all $\hat{\psi} \in W^{1,\infty}(\hat{K})$,

$$|\hat{E}(\hat{\psi}\hat{w})| \le \hat{C}\|\hat{E}(\cdot\ \hat{w})\|_{\mathcal{L}(W^{1,\infty}(\hat{K}))}|\hat{\psi}|_{W^{1,\infty}(\hat{K})} \le \hat{C}\|\hat{w}\|_{L^2(\hat{K})}|\hat{\psi}|_{W^{1,\infty}(\hat{K})}.$$

Take $\hat{\psi} = \hat{q}_0\hat{v}$. We check that

$$\begin{aligned} |\hat{\psi}|_{W^{1,\infty}(\hat{K})} &\le |\hat{q}_0|_{W^{1,\infty}(\hat{K})}\|\hat{v}\|_{L^\infty(\hat{K})} + \|\hat{q}_0\|_{L^\infty(\hat{K})}|\hat{v}|_{W^{1,\infty}(\hat{K})} \\ &\le \hat{C}\left(|\hat{q}_0|_{W^{1,\infty}(\hat{K})}\|\hat{v}\|_{L^2(\hat{K})} + \|\hat{q}_0\|_{L^\infty(\hat{K})}|\hat{v}|_{H^1(\hat{K})} \right). \end{aligned}$$

The last inequality holds because $\hat{v} \in P_1(\hat{K})$ and all norms on $P_1(\hat{K})$ are equivalent. Finally, recall the relations [12, Theorem 3.1.2] that for any integer $m \ge 0$, any $q \in [1, \infty]$, and for any $\psi \in W^{m,p}(K)$,

$$|\hat{\psi}|_{W^{m,q}(\hat{K})} \le C\|B\|^m |\det(B)|^{-\frac{1}{q}}|v|_{W^{m,q}(K)}.$$

Combine all the estimates above, we obtain the desired estimate. □

For the heterogeneous multi-scale error, we have the following result. We do not intend to make these estimates sharp.

Theorem 3.3. *Let dimension $d = 2$ and $\varepsilon \ll \delta \lesssim h$. Let $e(\mathrm{HMS})$ be the multi-scale heterogeneous error defined in (3.2), we have, for $\varepsilon \geq 0$ sufficiently small, the following estimate:*

$$\mathbb{E}\; e(\mathrm{HMS}) \leq \begin{cases} C\left(\dfrac{\varepsilon}{\delta}\right)^{\frac{d}{2}}, & \text{in the SRC setting,} \\ C\left(\dfrac{\varepsilon}{\delta}\right)^{\frac{\alpha}{2}}, & \text{in the LRC setting.} \end{cases} \tag{3.5}$$

The constants C above does not depend on h, δ or ε.

Proof. For any $v^h \in V^h$, its restriction $v^h|_K$ is in $P_1(K)$. Therefore, according to the definition of $e(\mathrm{HMS})$ in (3.2), we try to estimate

$$e_K := \sup_{v,w \in P_1(K)} \frac{|K||\hat{\mathcal{A}}_\varepsilon^{h,\delta}(u^h, v^h)[x_K] - \hat{\mathcal{A}}_0^{h,\delta}(u^h, v^h)[x_K]|}{\|v\|_{H^1(K)}\|w\|_{H^1(K)}}.$$

Since $\hat{\mathcal{A}}_\varepsilon^{h,\delta} - \hat{\mathcal{A}}_0^{h,\delta}$ may be viewed as a bilinear form, we have

$$e_K \leq \Big[\sum_{m,n=1}^{3} |K|^2 \big|\hat{\mathcal{A}}_\varepsilon^{h,\delta}(\phi_m, \phi_n)[x_K] - \hat{\mathcal{A}}_0^{h,\delta}(\phi_m, \phi_n)[x_K]\big|^2\Big]^{\frac{1}{2}},$$

where $\{\phi_m, m = 1, 2, 3\}$ are some orthonormal basis of $P_1(K)$. Let $X_{m,n}^\varepsilon$ denotes the difference $|K|\big(\hat{\mathcal{A}}_\varepsilon^{h,\delta}(\phi_m, \phi_n)[x_K] - \hat{\mathcal{A}}_0^{h,\delta}(\phi_m, \phi_n)[x_K]\big)$. By Cauchy-Schwarz,

$$\mathbb{E}\; e_K \leq \Big[\sum_{m,n=1}^{3} \mathbb{E}|X_{m,n}^\varepsilon|^2\Big]^{\frac{1}{2}}.$$

Therefore, we estimate $\mathbb{E}|X_{m,n}^\varepsilon|^2$. From the definition, we check that

$$X_{m,n}^\varepsilon = \int q_\varepsilon(x)\psi_{m,n}(x)dx, \quad \text{with} \quad \psi_{m,n}(x) = \chi_{K_\delta}(x)\frac{|K|}{|K_\delta|}\phi_m(x)\phi_n(x).$$

Here and in the sequel, χ_A is the indication function for a set $A \subset \mathbb{R}^2$. Abusing notations, we use X^ε and ψ instead of $X_{m,n}^\varepsilon$ and $\psi_{m,n}$ in the following.

Let us estimate $\mathbb{E}|X^\varepsilon|^2$. In the SRC setting, it is known that $\varepsilon^{-\frac{d}{2}}X^\varepsilon$ converges in distribution to a mean-zero Gaussian variable with variance $\sigma^2\|\psi\|_{L^2}^2$; see [5, Theorem 3.8]. Therefore, for sufficiently small ε, we have

$$\mathbb{E}|X_\varepsilon|^2 \leq C\varepsilon^d\|R\|_{L^1}\|\psi\|_{L^2}^2 = C\varepsilon^d\|R\|_{L^1}(\frac{h}{\delta})^{2d}\|\phi_m\phi_n\|_{L^2(K_\delta)}^2 \leq C\|R\|_{L^1}h^{2d}(\frac{\varepsilon}{\delta})^d. \tag{3.6}$$

Here R is the correlation function of q defined in (2.1). We used the fact that $|K|/|K_\delta| = (h/\delta)^d$ and we calculated that $\|\phi_m\phi_n\|_{L^2(K_\delta)}^2 \leq |K_\delta| \leq C\delta^d$.

In the LRC setting, it is known that $\varepsilon^{-\frac{\alpha}{2}}X^\varepsilon$ converges in distribution to a mean-zero Gaussian variable with variance $\|\psi\otimes\psi\|_{L^1(Y^2,\kappa|x-y|^{-\alpha}dxdy)}$; see [6, Lemma 4.3]. Consequently, for sufficiently small ε, we have

$$\begin{aligned}\mathbb{E}|X_\varepsilon|^2 &\le C\varepsilon^\alpha \iint_{K_\delta\times K_\delta} \frac{\kappa\psi(x)\psi(y)}{|x-y|^\alpha}dxdy \le C\varepsilon^\alpha \|\psi\|^2_{L^{\frac{4}{4-\alpha}}}\\ &= C\varepsilon^\alpha(\frac{h}{\delta})^{2d}\|\phi_m\phi_n\|^2_{L^{\frac{4}{4-\alpha}}} \le Ch^{2d}(\frac{\varepsilon}{\delta})^\alpha.\end{aligned} \tag{3.7}$$

In the second inequality we used Hardy-Littlewood-Sobolev inequality [22, Theorem 4.3], and we calculated that $\|\phi_m\phi_n\|_{L^{\frac{4}{4-\alpha}}} \le |K_\delta|^{\frac{4-\alpha}{4}}$.

The inequalities (3.6) and (3.7) show that $\mathbb{E}\ e_K$ is of order $h^d(\frac{\varepsilon}{\delta})^{d/2}$ and $h^d(\frac{\varepsilon}{\delta})^{\alpha/2}$ in the SRC and LRC settings respectively. From the definition of $e(\text{HMS})$, we see that

$$\mathbb{E}\ e(\text{HMS}) \le \sum_{K\in\mathcal{T}_h} \mathbb{E}\ e_K \le \frac{2}{h^2}\sup_{K\in\mathcal{T}_h} \mathbb{E}\ e_K. \tag{3.8}$$

Here, $\frac{2}{h^2}$ is a bound for the total number of elements in $\mathcal{T}_h$. Since the estimates (3.6) and (3.7) are uniform over $K\in\mathcal{T}_h$, we obtain the desired estimates. $\square$

3.2 Coordinate representation and conservative form

The next step is to reformulate the multi-scale schemes (1.8) and (1.10) as linear systems for the coordinates of the solutions in V^h, to investigate the structure of the associated stiffness matrices, and to write the discrete corrector $u_\varepsilon^{h,\delta} - u_0^{h,\delta}$ in terms of their coordinates.

We start by introducing some useful notation. In the triangulation illustrated by Fig. 3, we identify each grid point (ih, jh) with a unique two dimensional index (i,j). The set of inner grid points are denoted by $\mathcal{I} = \{(i,j) \mid 1 \le i,j \le N-1\}$, and the set of all grid points including the boundary ones is denoted by $\overline{\mathcal{I}} = \{(i,j) \mid 0 \le i,j \le N\}$. We define six difference operators $\mathrm{d}_s^\pm : \mathcal{I} \to \overline{\mathcal{I}}$ as follows:

$$\mathrm{d}_1^\pm(i,j) = (i\pm 1, j), \quad \mathrm{d}_2^\pm(i,j) = (i, j\pm 1), \quad \mathrm{d}_3^\pm(i,j) = (i\pm 1, j\pm 1). \tag{3.9}$$

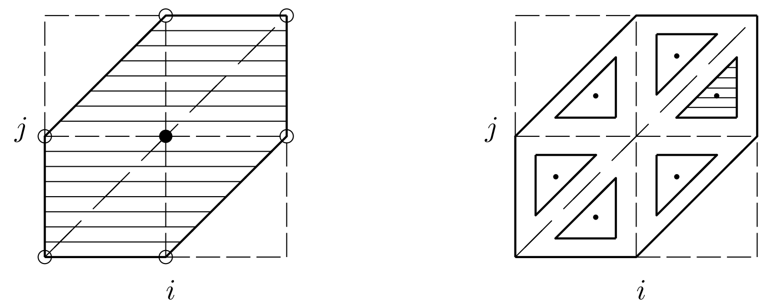
Here, $s = 1,2,3$ denotes three directions: horizontal, vertical and diagonal; the plus or minus sign indicates forward or backward differences.

In the sequel, we often write (i,j) simply as ij. For each $ij\in\mathcal{I}$, there corresponds a basis function ϕ^{ij} which is piecewise linear on each element $K\in\mathcal{T}_h$, has value one at ij and has value zero at other nodal points. Any function v^h in the space V^h can be uniquely written as $v^h(x) = \sum_{ij\in\mathcal{I}} V_{ij}\phi^{ij}(x)$, and the vector $(V_{ij}) \in \mathbb{R}^{(N-1)\times(N-1)}$ is called the coordinates of v^h. We identify $\mathbb{R}^{(N-1)\times(N-1)}$, the space for the coordinates, with V^h itself. Now, the difference operators $\mathrm{d}_s^\pm$ induce difference operators $D_s^\pm$ on V^h as follows:

$$D_s^+ V_{ij} = V_{\mathrm{d}_s^+ ij} - V_{ij}, \quad D_s^- V_{ij} = V_{ij} - V_{\mathrm{d}_s^- ij}. \tag{3.10}$$

Note when $\mathrm{d}_s^\pm ij$ lands outside of $\mathcal{I}$, *i.e.* on the boundary, the value $V_{\mathrm{d}_s^\pm ij}$ is set to zero.

Figure 4: Left: The support of a basis function ϕ^{ij}, denoted by $\mathcal{K}_{ij}$. Right: The shrunk integral region $\mathcal{K}_{ij}^{\delta}$.



Using the coordinate representation of functions $u_\varepsilon^{h,\delta} = \sum_{ij} U_{ij}^{\varepsilon} \phi^{ij}$ and $u_0^{h,\delta} = \sum_{ij} U_{ij}^{0} \phi^{ij}$, we can recast the heterogeneous multi-scale schemes (1.8) and (1.10) as the following systems:

$$A^{\varepsilon}_{ijkl} U^{\varepsilon}_{kl} = \langle f, \phi^{ij} \rangle, \tag{3.11}$$

$$A^{0}_{ijkl} U^{0}_{kl} = \langle f, \phi^{ij} \rangle. \tag{3.12}$$

Here, the stiffness matrices are defined by

$$A^{\varepsilon}_{ijkl} = \mathcal{A}^{h,\delta}_{\varepsilon}(\phi^{ij}, \phi^{kl}), \quad A^{0}_{ijkl} = \mathcal{A}^{h,\delta}_{0}(\phi^{ij}, \phi^{kl}).$$

These stiffness matrices have the following structures.

Proposition 3.4. *Let $A^p = (A^p_{ijkl})$ with $p = 0$ or ε be the stiffness matrices above. We observe*

(P1) $A^p_{ijkl} = A^p_{klij}$;

(P2) $A^p_{ijkl} = 0$ *unless* $kl \in \mathcal{I}_{ij} := \{ij\} \bigcup \{\mathrm{d}_s^{\pm} ij \mid s = 1, 2, 3\}$.

(P3) *For any $ij \in \mathcal{I}$, we have*

$$A^p_{ijij} = d^p_{ij} - \sum_{s=1}^{3} \Big(A^p_{ij\mathrm{d}_s^+ ij} + A^p_{ij\mathrm{d}_s^- ij} \Big), \tag{3.13}$$

for some d^p_{ij} that can be explicitly computed as in (3.14) *below.*

Proof. The first two observations are obvious, so only the third one needs to be stressed. According to (1.8) and (1.10), to calculate A^p_{ijij} we need to integrate the function $|\nabla \phi^{ij}(x)|^2 + q^p(x)|\phi^{ij}(x)|^2$. We observe that the support of ϕ^{ij}, denoted by $\mathcal{K}_{ij}$, is a hexagon consisting of six triangle elements as illustrated in Fig. 4-Left. The integration is actually taken over $\mathcal{K}_{ij}^{\delta}$, the region obtained by shrinking the triangle elements in $\mathcal{K}_{ij}$ with respect to their barycenters as illustrated in Fig. 4-Right. Let us consider a typical triangle in $\mathcal{K}_{ij}$ with nodal points $ij, \mathrm{d}_1^+ ij$ and $\mathrm{d}_3^+ ij$. Abusing notation, we call it K and the corresponding smaller triangle K_δ. Note K_δ corresponds to the shaded region in the figure. On this region, the three non-zero basis functions are $\phi^{ij}, \phi^{\mathrm{d}_1^+ ij}$ and $\phi^{\mathrm{d}_3^+ ij}$. They satisfy

$$\phi^{ij} + \phi^{\mathrm{d}_1^+ ij} + \phi^{\mathrm{d}_3^+ ij} = 1, \quad \nabla\phi^{ij} + \nabla\phi^{\mathrm{d}_1^+ ij} + \nabla\phi^{\mathrm{d}_3^+ ij} = 0.$$

Multiply ϕ^{ij} on both sides of the first equation, and $\nabla\phi^{ij}$ on the second equation. We have

$$(\phi^{ij})^2 = \phi^{ij} - (\phi^{\mathrm{d}_1^+ ij} + \phi^{\mathrm{d}_3^+ ij})\phi^{ij}, \quad |\nabla\phi^{ij}|^2 = -(\nabla\phi^{\mathrm{d}_1^+ ij} + \nabla\phi^{\mathrm{d}_3^+ ij}) \cdot \nabla\phi^{ij}.$$

Consequently, we have

$$\begin{aligned}
\hat{\mathcal{A}}_p^{h,\delta}(\phi^{ij}, \phi^{ij})[x_K] &= ⨍_{K_\delta} |\nabla\phi^{ij}|^2 + q^p|\phi^{ij}|^2 dx \\
&= ⨍_{K_\delta} q^p \phi^{ij} dx - \sum_{s=1,3} ⨍_{K_\delta} \nabla\phi^{ij} \cdot \nabla\phi^{\mathrm{d}_s^+ ij} + q^p\phi^{ij}\phi^{\mathrm{d}_s^+ ij} dx \\
&= ⨍_{K_\delta} q^p \phi^{ij} dx - \hat{\mathcal{A}}_p^{h,\delta}(\phi^{ij}, \phi^{\mathrm{d}_1^+ ij})[x_K] - \hat{\mathcal{A}}_p^{h,\delta}(\phi^{ij}, \phi^{\mathrm{d}_s^+ ij})[x_K].
\end{aligned}$$

Summing over the integrals on all six triangles, and using the notations of A^p, $\mathcal{A}_p^{h,\delta}$ and $\hat{\mathcal{A}}_p^{h,\delta}$, $p = 0, \varepsilon$, we see that (3.13) holds with d_{ij}^p defined by

$$d_{ij}^p = \sum_{K \in \mathcal{K}_{ij}} |K| ⨍_{K_\delta} q^p \phi^{ij} dx. \tag{3.14}$$

This completes the proof. $\square$

It follows immediately that the matrix A^p acts on vectors in V^h as follows:

$$(A^p V)_{ij} = \sum_{s=1}^{3} D_s^+(\alpha_{ij}^{s,p} D_s^- V_{ij}) + d_{ij}^{\varepsilon} V_{ij},$$

where $\alpha_{ij}^{s,p}$ is short-hand notation for $A^p_{ij\mathrm{d}_s^- ij}$ and it has the expression

$$\alpha_{ij}^{s,p} := \sum_{K \in \mathcal{K}_{ij}} |K| ⨍_{K_\delta} \nabla\phi^{ij} \cdot \nabla\phi^{\mathrm{d}_s^- ij} + q^p \phi^{ij} \phi^{\mathrm{d}_s^- ij} dx.$$

Note that when $\mathrm{d}_s^{\pm} ij$ lands outside of $\mathcal{I}$, *i.e.* on the boundary, $\phi^{\mathrm{d}_s^{\pm} ij}$ is the unique continuous function which is linear on each $K \in \mathcal{T}_h$, has value one at $\mathrm{d}_s^{\pm} ij$ and value zero at all other nodal points. Finally, taking the difference of A^ε and A^0 we obtain

$$(A^\varepsilon V - A^0 V)_{ij} = \sum_{s=1}^{3} D_s^+(\alpha_{\varepsilon ij}^s D_s^- V_{ij}) + d_{\varepsilon ij} V_{ij}, \tag{3.15}$$

where the vectors $(\alpha^s_{\varepsilon ij})$ and $(d_{\varepsilon ij})$ are

$$\alpha^s_{\varepsilon ij} := \alpha^{s,\varepsilon}_{ij} - \alpha^{s,0}_{ij} = \sum_{K\in\mathcal{K}_{ij}} |K| \fint_{K_\delta} q_\varepsilon \phi^{ij}\phi^{\mathrm{d}_s^- ij} dx, \tag{3.16}$$

$$d_{\varepsilon ij} := d^\varepsilon_{ij} - d^0_{ij} = \sum_{K\in\mathcal{K}_{ij}} |K| \fint_{K_\delta} q_\varepsilon \phi^{ij} dx. \tag{3.17}$$

Formula (3.15) is essential in our analysis because it provides an explicit expression of the discrete corrector $u^{h,\delta}_\varepsilon - u^{h,\delta}_0$. Identify these solutions with the vectors (U^ε_{ij}) and (U^0_{ij}) in (3.11-3.12). We verify that

$$A^0_{ijkl}(U^\varepsilon - U^0)_{kl} = -(A^\varepsilon - A^0)_{ijkl} U^\varepsilon_{kl}.$$

Here and after, we will use the summation convention that repeated indices are summed over unless otherwise stated. Let (G^0_{ijkl}) be the inverse of A^0. We then have

$$(U^\varepsilon - U^0)_{ij} = -G^0_{ijkl}(A^\varepsilon - A^0)_{klst} U^\varepsilon_{st}. \tag{3.18}$$

Using the formula (3.15) and summation by parts, we obtain

$$\begin{aligned}(U^\varepsilon - U^0)_{ij} &= -G_{ijkl}(A^\varepsilon - A^0)_{klts} U^0_{ts} - G_{ijkl}(A^\varepsilon - A^0)_{klts}(U^\varepsilon - U^0)_{ts} \\ &= \sum_{s=1}^3 (D^-_s G_{ijkl})(\alpha^s_\varepsilon D^-_s U^0)_{kl} - G_{ijkl}(d_\varepsilon U^0)_{kl} \\ &+ \sum_{s=1}^3 (D^-_s G_{ijkl})(\alpha^s_\varepsilon D^-_s (U^\varepsilon - U^0))_{kl} - G_{ijkl}(d_\varepsilon (U^\varepsilon - U^0))_{kl}.\end{aligned} \tag{3.19}$$

Note that for two vectors of the same dimension, say d_ε and U^0 above, the notation $(d_\varepsilon U^0)$ is the vector obtained by multiplying the corresponding components. This decomposition formula will be the starting point of our analysis in the next section.

4 Proof of the Main Results

In this section, we prove Theorem 1.2 using the coordinate representation (3.19) of the discrete corrector.

From (3.16) and (3.17) we see that α^s_ε and d_ε may be seen as averages of fluctuations and hence are asymptotically small. In the decomposition formula (3.19), the first sum involves linear terms of these random processes, while the second sum involves product of α^s_ε or d_ε with $U^\varepsilon - U^0$. The second term hence is much smaller if we can control $U^\varepsilon - U^0$. This is done in the following lemma.

Lemma 4.1. *Let U^ε_{ij} and U^0_{ij} be the coordinates of the numerical solutions to the random and the deterministic equations* (1.5) *and* (1.6) *respectively. Suppose that there exist some constants $C > 0, \gamma_j \in \mathbb{R}, j = 1, \cdots, 4$ so that*

$$|D^-_s G_{ijkl}| \le Ch^{\gamma_1}, |D^-_s U^\varepsilon_{ij}| \le Ch^{\gamma_2}, |G_{ijkl}| \le Ch^{\gamma_3} \,\textit{and}\, |U^\varepsilon_{ij}| \le Ch^{\gamma_4} \tag{4.1}$$

for any $s=1,2,3$. Let $d=2$. The following holds.

(1) *If the random process q satisfies the SRC setting, we have*

$$\mathbb{E}\|U^\varepsilon - U^0\|_{\ell^2}^2 \le Ch^{2(\min\{\gamma_1+\gamma_2,\gamma_3+\gamma_4\})}\|R\|_1\left(\frac{\varepsilon}{\delta}\right)^d. \tag{4.2}$$

(2) *If the random process q satisfies the LRC setting, we have*

$$\mathbb{E}\|U^\varepsilon - U^0\|_{\ell^2}^2 \le C(\alpha,\kappa)h^{2(\min\{\gamma_1+\gamma_2,\gamma_3+\gamma_4\})}\left(\frac{\varepsilon}{\delta}\right)^\alpha. \tag{4.3}$$

The constant C does not depend on h,δ or ε.

Proof. We observe that (3.19) can also be written as

$$(U^\varepsilon - U^0)_{ij} = \sum_{s=1}^{3}(D_s^- G_{ij})_{kl}(\alpha_\varepsilon^s D_s^- U^\varepsilon)_{kl} - G_{ijkl}(d_\varepsilon U^\varepsilon)_{kl}.$$

Using the bounds in (4.1) and Cauchy-Schwarz, we have

$$\mathbb{E}|U^\varepsilon - U^0|_{ij}^2 \le CN^2h^{2(\gamma_1+\gamma_2)}\sum_{s=1}^{3}\sum_{kl}\mathbb{E}|\alpha_{\varepsilon kl}^s|^2 + CN^2h^{2(\gamma_3+\gamma_4)}\sum_{kl}\mathbb{E}|d_{\varepsilon kl}|^2. \tag{4.4}$$

Here, N^2 is the number of nodal points, *i.e.* $N^2 = |\mathcal{I}| \simeq h^{-2}$. Take expectation. It suffices to estimate $\mathbb{E}|\alpha_{\varepsilon kl}^s|^2$ and $\mathbb{E}|d_{\varepsilon kl}|^2$. We rewrite (3.16-3.17) as

$$\alpha_{\varepsilon kl}^s = \int q_\varepsilon(x)a_{kl}^s(x)dx, \qquad d_{\varepsilon kl} = \int q_\varepsilon(x)b_{kl}(x)dx, \tag{4.5}$$

with a_ε^s and b_ε defined by

$$a_{kl}^s(x) = \sum_{K\in\mathcal{K}_{kl}}\frac{|K|}{|K_\delta|}\chi_{K_\delta}(x)\phi^{kl}(x)\phi^{\mathrm{d}_s^- kl}(x), \quad b_{kl}(x) = \sum_{K\in\mathcal{K}_{kl}}\frac{|K|}{|K_\delta|}\chi_{K_\delta}(x)\phi^{kl}(x). \tag{4.6}$$

Note that $|K|/|K_\delta| = (h/\delta)^d$. Note also that a_{kl}^s and b_{kl} are uniformly bounded on Y. Hence, we recognize $\alpha_{\varepsilon kl}^s$ and $d_{\varepsilon kl}$ as oscillatory integrals of uniformly bounded functions against fast varying mean-zero random processes. Such integrals are well understood. In fact, α_ε^s has the same form as X^ε in the proof of Theorem 3.3 and can be estimated in the same manner. In the SRC setting, we have that

$$\mathbb{E}|\alpha_{\varepsilon kl}^s|^2 \le C\varepsilon^d\|R\|_{L^1}\|a_{kl}^s\|_{L^2}^2 \le C\|R\|_{L^1}h^{2d}(\frac{\varepsilon}{\delta})^d. \tag{4.7}$$

In the LRC setting, the above estimate should be changed to

$$\mathbb{E}|\alpha_{\varepsilon kl}^s|^2 \le C\varepsilon^\alpha\|a_{kl}^s\otimes a_{kl}^s\|_{L^1(Y\times Y,|x-y|^{-\alpha}dxdy)} \le C(\alpha,\kappa)h^{2d}(\frac{\varepsilon}{\delta})^\alpha. \tag{4.8}$$

The mean square of $d_{\varepsilon kl}$ can be similarly estimated. Substitute these estimates into (4.4) to control the mean square of $(U^\varepsilon - U^0)_{ij}$; note that the sum over kl introduces a factor of h^{-d} which is the number of items in the sum. The estimates of $(U^\varepsilon - U^0)_{ij}$ are uniform in ij, summation over ij yields the desired results. Note that this additional summation introduces another h^{-d} to the estimates. □

Lemma 4.2. *Under the same condition of the previous lemma, we have*

$$(U^\varepsilon - U^0)_{ij} = \sum_{s=1}^{3} (D_s^- G_{ijkl})(\alpha_\varepsilon^s D_s^- U^0)_{kl} + G_{ijkl}(d_\varepsilon U^0)_{kl} + r_{ij}^\varepsilon. \tag{4.9}$$

Further, the error term r_{ij}^ε satisfies

$$\sup_{ij} \mathbb{E}|r_{ij}^\varepsilon| \le \begin{cases} Ch^{\min\{\gamma_1,\gamma_3\}+\min\{\gamma_1+\gamma_2,\gamma_3+\gamma_4\}+\frac{d}{2}} \left(\frac{\varepsilon}{\delta}\right)^d, & \text{in the SRC setting,} \\ Ch^{\min\{\gamma_1,\gamma_3\}+\min\{\gamma_1+\gamma_2,\gamma_3+\gamma_4\}+\frac{d}{2}} \left(\frac{\varepsilon}{\delta}\right)^\alpha, & \text{in the LRC setting.} \end{cases} \tag{4.10}$$

Proof. The decomposition holds with

$$r_{ij}^\varepsilon = \sum_{s=1}^{3} \sum_{kl} (D_s^- G_{ijkl}) \alpha_{\varepsilon kl}^s D_s^- (U_{kl}^\varepsilon - U^0)_{kl} + \sum_{kl} G_{ijkl} d_{\varepsilon kl} (U^\varepsilon - U^0)_{kl}. \tag{4.11}$$

Bound the $D_s^- G_{ijkl}$ and G_{ijkl} terms by (4.1), and use Cauchy-Schwarz. We get

$$|r_{ij}^\varepsilon| \le Ch^{\gamma_1} \sum_{s=1}^{3} \|\alpha_\varepsilon^s\|_{\ell^2} \|D_s^-(U^\varepsilon - U^0)\|_{\ell^2} + Ch^{\gamma_3} \|d_\varepsilon\|_{\ell^2} \|U^\varepsilon - U^0\|_{\ell^2}.$$

Note that $\|D_s^-(U^\varepsilon - U^0)\|_{\ell^2}^2 \le C\|U^\varepsilon - U^0\|_{\ell^2}$. Take expectation and use Cauchy-Schwarz again to get

$$\mathbb{E}|r_{ij}^\varepsilon| \le Ch^{\gamma_1} \sum_{s=1}^{3} \left(\mathbb{E}\|\alpha_\varepsilon^s\|_{\ell^2}^2 \mathbb{E}\|U^\varepsilon - U^0\|_{\ell^2}^2\right)^{\frac{1}{2}} + Ch^{\gamma_3} \left(\mathbb{E}\|d_\varepsilon\|_{\ell^2}^2 \mathbb{E}\|U^\varepsilon - U^0\|_{\ell^2}^2\right)^{\frac{1}{2}}. \tag{4.12}$$

Summing over kl in the estimates (4.7) and (4.8), we have

$$\mathbb{E}\|\alpha_\varepsilon^s\|_{\ell^2}^2 \le \begin{cases} C\|R\|_{L^1} h^d (\frac{\varepsilon}{\delta})^d, & \text{in the SRC setting,} \\ C(\alpha,\kappa) h^d (\frac{\varepsilon}{\delta})^\alpha, & \text{in the LRC setting.} \end{cases}$$

The same estimates hold also for $\mathbb{E}\|d_\varepsilon\|_{\ell^2}^2$. Substituting these estimates, together with (4.2-4.3), into (4.12) completes the proof. □

Remark 4.3. The assumption (4.1) is not a restriction because γ_j there can be negative. Indeed, consider an arbitrary triangle element K. Without loss of generality, let its vertices be $\{ij, i-1j, ij+1\}$. For any function $v^h \in V^h$ with coordinate vector V, we verify that

$$C_1 h^2 (|V_{ij}|^2 + |V_{i-1j}|^2 + |V_{ij+1}|^2) \le \|v^h\|_{L^2(K)}^2 \le C_2 h^2 (|V_{ij}|^2 + |V_{i-1j}|^2 + |V_{ij+1}|^2). \tag{4.13}$$

This is due to the fact that all norms on the finite dimensional space $V^h|_K$ are equivalent and h^2 is the right scaling. We can check also that $\nabla v^h|_K$ is a constant vector $(D_1^- V_{ij}, D_2^- V_{ij+1})/h$. It follows that

$$|v|_{W^{1,q}(K)} = Ch^{\frac{2}{q}-1} \|(D_1^- V_{ij}, D_2^- V_{ij+1})\|_q = Ch^{\frac{2}{q}-1} \left(|D_1^- V_{ij}|^q + |D_2^- V_{ij+1}|^q\right)^{\frac{1}{q}}. \tag{4.14}$$

Now for $u_\varepsilon^{h,\delta}$, we know its H^1 norm is bounded independent of h and ε. Applying the results above we find that $|U_{ij}^\varepsilon| \le Ch^{-1}$ and $|D_1^- U_{ij}^\varepsilon| \le C$. Other elements of U^ε and $D_s^- U^\varepsilon$ can be estimated in the same way. Hence, we may choose $\gamma_2 = 0$ and $\gamma_4 = -1$. Similarly, the discrete Green's function $G^h(x,y) = G_{ijkl}\phi^{ij}(x)\phi^{kl}(y)$ is known to have $W^{1,q}$ norm for some $q < 2$ bounded by $C|\log h|$ for any fixed x; see [17, Theorem 5.1]. Using (4.13) and (4.14) we may choose γ_1 and γ_3 properly, say $\gamma_1 = \gamma_3 = -1$. $\square$

Now we are ready to prove the main theorem. We denote the solution operator of the averaged scheme (1.10) by $\mathcal{G}^{h,\delta}$. That is, $\mathcal{G}^{h,\delta} f$ is the unique solution in V^h that satisfies (1.10). Using the coordinate representation and summation convention, we verify that

$$\mathcal{G}^{h,\delta} f(x) = G_{ijkl}\langle f, \phi^{kl}\rangle \phi^{ij}(x). \tag{4.15}$$

Proof of Theorem 1.2. Take any test function $\varphi \in \mathcal{C}^2(\overline{Y})$. Let us denote the function $\mathcal{G}^{h,\delta}\varphi$ by m^h and its coordinate vector by (M_{ij}). Let $\beta = d$ in the SRC setting and $\beta = \alpha$ in the LRC setting. From the decomposition (4.9), we write

$$\begin{aligned}
&\frac{1}{\sqrt{\varepsilon^\beta}}\int_Y \varphi(x)[u_\varepsilon^{h,\delta} - u_0^{h,\delta}]dx = \frac{1}{\sqrt{\varepsilon^\beta}}(U^\varepsilon - U^0)_{ij}\langle \varphi, \phi^{ij}\rangle \\
=&\frac{1}{\sqrt{\varepsilon^\beta}}\left[\left(\sum_{s=1}^3 D_s^- G_{ijkl}(\alpha_\varepsilon^s D_s^- U^0)_{kl} + G_{ijkl}(d_\varepsilon U^0)_{kl} + r_{ij}^\varepsilon\right)\langle \varphi, \phi^{ij}\rangle\right] \\
=&\frac{1}{\sqrt{\varepsilon^\beta}}\left[\sum_{s=1}^3 D_s^- M_{kl}(\alpha_\varepsilon^s D_s^- U^0)_{kl} + M_{kl}(d_\varepsilon U^0)_{kl}\right] + \frac{1}{\sqrt{\varepsilon^\beta}} r_{ij}^\varepsilon \langle \varphi, \phi^{ij}\rangle.
\end{aligned} \tag{4.16}$$

In the last equality, we used the fact that $G_{ijkl} = G_{klij}$ and recognized the coordinate M_{kl} according to the formula (4.15).

First convergence as $\varepsilon \to 0$ while h is fixed. Let us control the last term first. Thanks to the estimate (4.10), we have

$$\mathbb{E}\left|\frac{1}{\sqrt{\varepsilon^\beta}} r_{ij}^\varepsilon\langle\varphi,\phi^{ij}\rangle\right| \le \frac{1}{\sqrt{\varepsilon^\beta}}\big(\mathbb{E}|r_{ij}^\varepsilon|\big)|\langle \varphi, \phi^{ij}\rangle| \le C(h)\|\varphi\|_{L^1}\sqrt{\varepsilon^\beta}. \tag{4.17}$$

This term hence converges to zero in $L^1(\mathbb{P})$ and does not contribute to the limiting distribution. So we focus on the other two terms that are linear in α_ε^s and d^ε, respectively. Substituting the oscillatory integral representations (4.6) into (4.16), we find that

$$\begin{aligned}
\frac{1}{\sqrt{\varepsilon^\beta}}\int_Y \varphi(x)[u_\varepsilon^{h,\delta} - u_0^{h,\delta}]dx &\simeq \frac{1}{\sqrt{\varepsilon^\beta}}\int_Y q_\varepsilon(x) L_1^{h,\delta}(x)dx + \frac{1}{\sqrt{\varepsilon^\beta}}\int_Y q_\varepsilon(x)L_2^{h,\delta}(x)dx \\
&= \frac{1}{\sqrt{\varepsilon^\beta}}\int_Y q_\varepsilon(x) L^{h,\delta}(x)dx.
\end{aligned} \tag{4.18}$$

Here, $L_j^{h,\delta}$, $j = 1, 2$ and $L^{h,\delta} = L_1^{h,\delta} + L_2^{h,\delta}$ depend on φ through M and are defined by

$$
\begin{aligned}
L_1^{h,\delta}(x) &= \sum_{s=1}^{3} D_s^- M_{kl} (a_{kl}^s(x) D_s^- U^0)_{kl} \\
&= \sum_{kl} \sum_{K \in \mathcal{K}_{kl}} \frac{|K|}{|K_\delta|} \chi_{K_\delta}(x) \sum_{s=1}^{3} (D_s^- M_{kl})(D_s^- U_{kl}^0) \phi^{kl}(x) \phi^{\mathrm{d}_s^- kl}(x), \\
L_2^{h,\delta}(x) &= b_{kl}(x)(MU^0)_{kl} = \sum_{kl} \sum_{K \in \mathcal{K}_{kl}} \frac{|K|}{|K_\delta|} \chi_{K_\delta}(x) M_{kl} U_{kl}^0 \phi^{kl}(x) \\
&= \sum_{K \in \mathcal{T}_h} \frac{|K|}{|K_\delta|} \chi_{K_\delta}(x) \sum_{kl \in \mathcal{I}_K} M_{kl} U_{kl}^0 \phi^{kl}(x) = \sum_{K \in \mathcal{T}_h} \frac{|K|}{|K_\delta|} \chi_{K_\delta}(x) \varPi^h(m^h u_0^{h,\delta})(x).
\end{aligned} \tag{4.19}
$$

Here, $\mathcal{I}_K$ contains the indices so that (kh, lh) is in $\overline{K}$ and $\varPi^h(m^h u_0^{h,\delta})$ is the projection in V^h of the function $m^h u_0^{h,\delta}$. Now the convergence in item one and the first conclusion of item three of Theorem (1.2) follows from the representation (4.18) and the well-known results on limiting distribution of oscillatory integrals; we refer the reader to Theorem 3.8 of [5] for the SRC setting, and to Lemma 4.3 of [6] for the LRC setting.

Second convergence as $h \to 0$, SRC setting. Now we prove item two of the theorem. It concerns the limiting distribution, as h goes to zero, of the Gaussian random variable which is obtained as the limiting distribution in the first step. This step depends on the correlation length of the random field and needs to be considered separately for the SRC and LRC setting. Here we focus on proving (1.12) first.

We have the following key observation:

$$
L_1^{h,\delta} \longrightarrow 0 \text{ in } L^\infty(Y) \text{ as } h \to 0. \tag{4.20}
$$

Indeed, for any fixed $x \in Y$, since $|\phi^{ij}| \le 1$ uniformly and $|K|/|K_\delta| = (h\delta^{-1})^2$, we have

$$
|L_1^{h,\delta}(x)| \le C \left(\frac{h}{\delta}\right)^2 h^2 \sum_{s=1}^{3} \left\| \frac{D_s^- M_{kl}}{h} \right\|_{\ell^2} \left\| \frac{D_s^- U_{kl}^0}{h} \right\|_{\ell^2} \le C \left(\frac{h}{\delta}\right)^2 h^2 |m^h|_{H^1} |u_0^{h,\delta}|_{H^1}.
$$

Since $u_0^{h,\delta}$ and m^h are yielded form the scheme (1.10) for smooth right hand side f and φ, they have bounded H^1 norms. We assume that the ratio h/δ is fixed while h is sent to zero. Therefore, the above estimate shows that $L_1^{h,\delta}$ goes to zero uniformly, proving the claim.

According to (4.18), the left hand side of (1.12) can be written as

$$
\sigma \int_Y L_1^{h,\delta}(x) dW(x) + \sigma \int_Y L_2^{h,\delta}(x) dW(x). \tag{4.21}
$$

Our plan is to show that the second term above converges to the right hand side of (1.12) while the first term above converges in probability to zero; this is indeed sufficient for (1.12). Since all random variables involved are Gaussian, we only need to calculate their variances. Thanks to Itô's isometry, we have

$$
\mathrm{Var}\, \sigma \int_Y L_1^{h,\delta}(x) dW(x) = \sigma^2 \int_Y |L_1^{h,\delta}(x)|^2 dx.
$$

Due to (4.20), the above variance goes to zero, proving our claim for the first term. For the second one, we have again

$$\mathrm{Var}\ \sigma \int_Y L_2^{h,\delta}(x)dW(x) = \sigma^2 \int_Y |L_2^{h,\delta}(x)|^2 dx = \left(\frac{\sigma h}{\delta}\right)^2 \sum_{K\in\mathcal{T}_h} |K| \fint_{K_\delta} |\Pi^h(m^h u_0^{h,\delta})(x)|^2 dx.$$

We recognize the sum in the last term as a barycenter approximation of the integral that gives the L^2 norm square of $\Pi^h(m^h u_0^{h,\delta})$. Thanks to Lemma 4.4 below, $\|\Pi^h(m^h u_0^{h,\delta})\|_{L^2}$ converges to $\|u_0\mathcal{G}\varphi\|_{L^2}$. This implies that the variance of the second term in (4.21) converges to $(\sigma h/\delta)^2\|u_0\mathcal{G}\varphi\|_{L^2}^2$, proving (1.12).

Second convergence as $h \to 0$, LRC setting. Now we prove (1.14). Like in (4.21), we can write the left hand side of (1.14) as a sum of two Gaussian random variables. Using a modified isometry, we write the variance of the first variable as

$$\mathrm{Var}\ \sigma \int_Y L_1^{h,\delta}(x)W^\alpha(dx) = \iint_{Y^2} \frac{\kappa L_1^{h,\delta}(x)L_1^{h,\delta}(y)}{|x-y|^\alpha} dxdy = \mathscr{I}(L_1^{h,\delta}).$$

Here, we define the operator $\mathscr{I} : L^{\frac{4}{4-\alpha}} \to \mathbb{R}$ as

$$\mathscr{I}(g) := \|g\otimes g\|_{L^1(Y^2,\kappa|x-y|^{-\alpha}dxdy)} = \iint_{Y^2} \frac{\kappa g(x)g(y)}{|x-y|^\alpha} dxdy. \tag{4.22}$$

Recalling the Hardy-Littlewood-Sobolev inequality, Theorem 4.3 of [22], we have

$$|\mathscr{I}(g)| \le \kappa C(\alpha)\|g\|^2_{L^{\frac{4}{4-\alpha}}}. \tag{4.23}$$

Due to (4.20), this term goes to zero again and does not contribute to the limiting distribution. For the contribution of $L_2^{h,\delta}$, we have

$$\begin{aligned}\mathrm{Var}\ \sigma \int_Y L_2^{h,\delta}(x)W^\alpha(dx) &= \iint_{Y^2} \frac{\kappa L_2^{h,\delta}(x)L_2^{h,\delta}(y)}{|x-y|^\alpha} dxdy \\ &= \sum_{K\in\mathcal{T}_h}\sum_{K'\in\mathcal{T}_h} |K|^2 \fint_{K_\delta}\fint_{K'_\delta} \frac{\kappa \Pi^h(m^h u_0^{h,\delta})(x)\Pi^h(m^h u_0^{h,\delta})(y)}{|x-y|^\alpha} dxdy.\end{aligned}$$

We recognize the last sum as the barycenter approximation of $\mathscr{I}(\Pi^h(m^h u_0^{h,\delta}))$. Now (4.23) shows that $\mathscr{I}$ is continuous on $L^{\frac{4}{4-\alpha}}$. Since $\alpha < 2$ and $\frac{4}{4-\alpha} < 2$, we have the inclusion $L^2(Y) \subset L^{\frac{4}{4-\alpha}}(Y)$. Therefore $\mathscr{I}$ is also continuous on $L^2(Y)$. Applying (4.24) with $f_1 = \varphi$ and $f_2 = f$, we conclude that $\mathscr{I}(\Pi^h(m^h u_0^{h,\delta}))$ converges to $\mathscr{I}(u_0\mathcal{G}\varphi)$. This proves (1.14) and completes the proof of the theorem. $\square$

It remains to prove the following key lemma concerning the convergence of product of solutions yielded from the averaged heterogeneous multi-scale scheme (1.10).

Lemma 4.4. *Let $\mathcal{G}^{h,\delta}$ be the Green's operator of the scheme* (1.10)*. For any two functions $f_j \in \mathcal{C}^2(\overline{Y})$, $j = 1, 2$, let $\Pi^h(\mathcal{G}^{h,\delta} f_1 \mathcal{G}^{h,\delta} f_2)$ be the projection in V^h of the product of $\mathcal{G}^{h,\delta} f_1$ and $\mathcal{G}^{h,\delta} f_2$. We have that*

$$\Pi^h(\mathcal{G}^{h,\delta} f_1 \mathcal{G}^{h,\delta} f_2) \xrightarrow{L^2} \mathcal{G} f_1 \mathcal{G} f_2, \qquad \text{as } h \to 0. \tag{4.24}$$

As before, $\mathcal{G}$ above is the Green's operator of the homogenized equation (1.6)*.*

Proof. To simplify notation, let us denote the function $\mathcal{G}^{h,\delta} f_j$ by $\tilde{u}_j^h$, the functions $\mathcal{G} f_j$ by u_j, $j = 1, 2$.

The key to the proof relies on L^∞ error estimates for finite element methods. Such results are classic for the scheme with $h = \delta$ as proved in [26, 28]. For $\delta < h$, as explained before we may view the scheme as the standard finite element with (barycenter) numerical integrations. L^∞ error estimates for such practical schemes are more involved but were obtained in [30, 17]. In particular, the piecewise linear FEM with numerical quadrature was considered in Theorem 5.1 of [17], and it was shown that

$$\|\tilde{u}_j^h - u_j\|_{L^\infty} \le Ch^2 |\log h| \|f_j\|_{W^{2,\infty}}.$$

Since $\tilde{u}_j^h$, $j = 1, 2$, are bounded, the above also implies that

$$\|\tilde{u}_1^h \tilde{u}_2^h - u_1 u_2\|_{L^\infty} \le Ch^2 |\log h| \|f_j\|_{W^{2,\infty}}^2. \tag{4.25}$$

In fact, Theorem 5.1 of [17] also shows that

$$\|\tilde{u}_j^h\|_{W^{1,\infty}} \le \|u_j^h\|_{W^{1,\infty}} + Ch|\log h|(\|u_j\|_{W^{2,\infty}} + \|f_j\|_{W^{2,\infty}}).$$

Here, u_j^h is the FEM solution with $h = \delta$. The above estimate shows that $\tilde{u}_j^h$ is in W^1_∞. Since u_j are bounded, we check that $\tilde{u}_1^h \tilde{u}_2^h \in W^1_\infty$. From classical interpolation estimates, e.g. taking $k = m = 0$, $p = \infty$ and $q = 2$ in Theorem 3.1.6 of [12], we have

$$\|\tilde{u}_1^h \tilde{u}_2^h - \Pi_K^h(\tilde{u}_1^h \tilde{u}_2^h)\|_{L^2(K)} \le C|K|^{\frac{1}{2}} h |\tilde{u}_1^h \tilde{u}_2^h|_{W^{1,\infty}(Y)}.$$

Here, Π_K^h is the projection on the triangle element K. Summing over $K \in \mathcal{T}_h$, we have

$$\|\tilde{u}_1^h \tilde{u}_2^h - \Pi^h(\tilde{u}_1^h \tilde{u}_2^h)\|_{L^2(Y)} \le Ch \|\tilde{u}_1^h \tilde{u}_2^h\|_{W^{1,\infty}}. \tag{4.26}$$

Note that (4.25) controls $\|\tilde{u}_1^h \tilde{u}_2^h\|_{W^{1,\infty}}$. Sending h to zero, we finish the proof. $\square$

References

[1] G. Allaire and R. Brizzi, *A multiscale finite element method for numerical homogenization*, Multiscale Model. Simul., 4 (2005), pp. 790–812 (electronic).

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

[3] I. Babuska, *Homogenization and its applications, mathematical and computational problems*, Numerical Solutions of Partial Differential Equations-III (SYNSPADE 1975, College Park, MD, May 1975) (B. Hubbard ed.), Academic Press, New York, 1976, pp. 89–116.

[4] I. Babuska, *Solution of interface by homogenization*. I,II,III, SIAM J. Math. Anal. 7 (1976), pp. 603–634, 635–645, 8 (1977), pp. 923–937.

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

[6] G. Bal, J. Garnier, Y. Gu, and W. Jing, *Corrector theory for elliptic equations with oscillatory and random potentials with long range correlations*, Asymptotic Analysis, 77 (2012), pp. 123–145.

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

[8] G. Bal and W. Jing, *Corrector theory for MsFEM and HMM in random media*, Multiscale Model. Simul., 9 (2011), pp. 1549–1587.

[9] G. Bal and K. Ren, *Physics-based models for measurement correlations: application to an inverse Sturm-Liouville problem*, Inverse Problems, 25 (2009), pp. 055006, 13.

[10] L. Berlyand and H. Owhadi, *Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast*, Arch. Ration. Mech. Anal., 198 (2010), no. 2, pp. 677–721.

[11] 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.

[12] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.

[13] P. Doukhan, *Mixing*, vol. 85 of Lecture Notes in Statistics, Springer-Verlag, New York, 1994. Properties and examples.

[14] 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.

[15] 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 (electronic).

[16] 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.

[17] C. I. Goldstein, *Variational crimes and L^∞ error estimates in the finite element method*, Math. Comp., 35 (1980), pp. 1131–1157.

[18] 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.

[19] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of differential operators and integral functionals*, Springer-Verlag, New York, 1994.

[20] D. Khoshnevisan, *Multiparameter processes*, Springer Monographs in Mathematics, Springer-Verlag, New York, 2002. An introduction to random fields.

[21] S. M. Kozlov, *The averaging of random operators*, Mat. Sb. (N.S.), 109(151) (1979), pp. 188–202, 327.

[22] E. H. Lieb and M. Loss, *Analysis*, vol. 14 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2001.

[23] J. Nolen and G. Papanicolaou, *Fine scale uncertainty in parameter estimation for elliptic equations*, Inverse Problems, 25 (2009), pp. 115021, 22.

[24] H. Owhadi and L. Zhang, *Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast*, Multiscale Model. Simul. 9 (2011), no. 4, pp. 1373–1398.

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

[26] R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp., 38 (1982), pp. 437–445.

[27] M. Reed and B. Simon, *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975.

[28] R. Scott, *Optimal L^∞ estimates for the finite element method on irregular meshes*, Math. Comp., 30 (1976), pp. 681–697.

[29] M. S. Taqqu, *Convergence of integrated processes of arbitrary Hermite rank*, Z. Wahrsch. Verw. Gebiete, 50 (1979), pp. 53–83.

[30] L. B. Wahlbin, *Maximum norm error estimates in the finite element method with isoparametric quadratic elements and numerical integration*, RAIRO Anal. Numér., 12 (1978), pp. 173–202.